University of Calgary Rapid Vehicle-based Methane Emissions Mapping System (PoMELO) Single-Blind Testing Results from the Methane Emissions Technology Evaluation Center (METEC)

Prepared by:

Thomas Barchyn, MSc., P.Geo. (B.C.)
University of Calgary
Calgary, Alberta, Canada
tbarchyn@ucalgary.ca

Chris Hugenholtz, PhD
University of Calgary
Calgary, Alberta, Canada
chhugenh@ucalgary.ca

23 January 2020

Disclaimer: While reasonable effort has been made to ensure the accuracy, reliability, and completeness of the information presented here, this report is made available without representation for its use in any particular situation and with the understanding that the reader accepts full liability for application of the contents, regardless of any fault, omissions, or negligence of University of Calgary or Colorado State University METEC personnel.

Summary

This report presents results from single-blind experiments evaluating the University of Calgary Rapid Vehicle-based Methane Emissions Mapping System (hereafter PoMELO), conducted at the Colorado State University Methane Emissions Technology Evaluation Center (hereafter METEC).

PoMELO is a system for mapping emissions points on upstream oil and gas pads. The maps of emitting equipment are then used to guide close-range inspections, usually while onsite. Testing was designed to evaluate the ability for the PoMELO system to correctly identify both emitting and not emitting equipment.

Mock upstream pads were configured with 0-6 emissions points and the PoMELO system was used to detect and map the emissions locations at the equipment group and equipment unit scale. Over 5 days of testing, 105 individual experiment pads were surveyed. Individual units of equipment consist of wellheads, tanks, and separators. Equipment groups consist of two or more units of equipment in close proximity. The PoMELO team was unaware of emissions configurations until after all experiments were finished, ensuring that the PoMELO team executed surveys blind.

At the equipment scale, the total sum emissions rates detected by PoMELO was 4.75 g/s, with 0.198 g/s missed (96.0 % of sum emissions rates detected). At the equipment group scale the total sum emissions rates detected by PoMELO was 4.90 g/s, with 0.0541 g/s missed (98.9 % of sum emissions rates detected). At the equipment unit scale, 95.6% of emitting equipment were correctly detected for close-range follow-up, but 47.1% of clean equipment were also detected. At the equipment group scale 98.5% of emitting equipment groups were detected with 16.7% of clean equipment groups erroneously detected.

The minimum equipment group emissions rate that was detected was 0.00158 ± 8.29 x 10^{-5} g/s (0.301 ± 0.0159 scfh CH_4, 0.201 ± 0.0106 m^3/day CH_4). Detection frequency was near to 100% across the full range of emissions rates tested, limiting inference about the causes of non-detections. It is likely that the PoMELO minimum detectable emissions rate is below the minimum controllable flow rate at METEC.

Results suggest the PoMELO system is very sensitive within the context of detecting emissions sources on upstream oil and gas pads, and attributing emissions sources to equipment units; pairing PoMELO with a close-range follow-up method could provide near equivalent emissions reductions when compared against a close-range only survey mode. However, there is room for enhancements in system efficiency to reduce unnecessary follow-up of non-emitting equipment.

Contents

Introduction

In the upstream oil and gas industry, hundreds of thousands of wellpads are distributed across different production areas. These pads typically contain one or several units of equipment such as wellheads, separators, tanks, risers, and other associated production equipment. Searching for anomalous emissions is an essential component of emissions management programs implemented by producers. Cumulatively, anomalous emissions can have a major effect on the carbon intensity of hydrocarbon production (Alvarez et al., 2018). As a result, leak detection and repair (LDAR) is now extensively mandated by regulators globally, and often part of voluntary emissions control programs by producers.

Traditional methods for LDAR involve time-consuming component-by-component inspections with either US Environmental Protection Agency Method 21 (M21) or Optical Gas Imaging (OGI) technologies. In response to the cost of these incumbent survey approaches, a variety of new methods have been developed (Fox et al., 2019a) or are under development. These new methods vary considerably in platform, sensitivity, and operations. Most are fundamentally quite different than M21 or OGI, which implement an exhaustive leak survey at the component-scale. Consequently, linking the performance of these 'alternative LDAR' methods to emissions reductions can be difficult (Fox et al., 2019b).

The University of Calgary PoMELO Padmapper system is an alternative LDAR technology and methodology designed to facilitate emissions and cost reductions. The system has several modes, but the mode tested here implements a strategy of flagging parts of a site that are emitting for follow-up with component-scale techniques. Following these flags, close-range inspections can then pinpoint the leaking component using traditional techniques and issue repair instructions, typically in the form of a leak tag. The strategy is to direct the application of time-intensive, close-range, component-wise techniques only to the portions of the pad that are emitting, saving time and cost.

The system consists of 3 instruments streaming data into a computer with customized algorithms: (i) high performance methane sensor, (ii) GNSS, and (iii) anemometer. The computer fuses instrument data streams at high frequency and uses these data in proprietary algorithms to automatically map the position of emissions sources on the pad.

Although PoMELO and similar technologies may offer promise, it is vital to quantify the effectiveness of the technology to understand what the system can do, base regulatory approval, and evaluate economics. Due to the differences between technologies, there is no standard set of metrics for evaluating technology and testing is often ad hoc and difficult to compare. Further, testing on real oil and gas pads is both difficult and lacks full control over emissions sources present.

In response to the difficulty of testing, and to develop a defensible and robust measurement of PoMELO effectiveness, experiments were conducted at the Colorado State University METEC facility in November 2019.

The METEC facility was developed to evaluate LDAR technologies designed for the upstream oil and gas industry. METEC consists of multiple mock oil and gas pads, located on the outskirts of Fort Collins, Colorado, USA (Figure 1). Real oil and gas equipment are installed with hidden leak points that are not known to the participants. There is no oil and gas production on the site, so the environment is completely controlled and there are no extraneous methane emissions. Each leak point can be configured to release a precisely metered volume of gas. Participants arrive at the pad blind to the real leak locations or rates, prepare a survey datasheet, and then compare results after the end of all experiments. Many surveys can be run in close succession, with controlled replicates in different conditions. METEC is the global-leading facility for LDAR technology testing. Further, the blind mock pad approach developed and implemented at METEC is the leading approach for generating useful data about LDAR technology effectiveness in upstream settings.

***Figure 1:** Mock equipment at the METEC site. The facility uses real production equipment that has been modified to have hidden leak points throughout. Participants cannot tell where emissions are located – making the facility a globally-leading test platform for upstream oil and gas LDAR technologies.*

This report details results from a set of experiments performed by the PoMELO system at METEC 18-22 November 2019. First, we further describe the PoMELO system and experiment protocol. Second, we examine detection and localization results submitted by the PoMELO team, comparing against real emissions locations as configured by the METEC team.

Terminology

We use the following terminology in this report. We clarify these commonly used terms to ensure there is no ambiguity:

Detection: This indicates an equipment unit or equipment group was identified by PoMELO as emitting and should be surveyed with a close-range follow-up method. To clarify, detections are properties of physical assets. We do not use the term detection to pertain to 'detection of plumes', or 'detections of methane anomalies' here. Although the system uses identification of plumes as part of its approach for detecting emitting equipment, these identifications are not useful unless they can be attributed to a piece of equipment. This definition of detection cannot be measured in a laboratory.

Equipment unit: A METEC coded unit of mock oil and gas equipment. Usually a specific tank, separator, or wellhead.

Equipment group: A METEC coded group of equipment. Equipment units at METEC are grouped into clusters of one to five similar equipment units.

Flag: Synonymous with detection.

Localization: This is the process of using measured data to locate an emissions source. We measure localization through detection statistics, rather than directly.

Tag: Repair instructions issued to repair crews. This can take the form of a physical tag attached to a mal-functioning piece of equipment with instructions. Tags are typically produced from close-range, component-wise survey methods; the PoMELO system in unable to produce tags.

PoMELO overview

The PoMELO system is designed to accelerate application of component-scale LDAR techniques (Figure 2). Although screening techniques can be quite sensitive (Fox et al., 2019a), most cannot pinpoint the exact component that is leaking and provide instructions for repair crews. This lack of component-scale specificity means any screening technology must be paired with conventional close-range inspection techniques to identify component sources. The PoMELO system is designed such that the equipment-scale screening, component-scale surveys, and possibly repair, are done in one visit by the same crew. This reduces communication overhead, delays, miscommunications, and helps resolve emitting equipment faster. In practice, the system provides instructions (termed ‘flags’) on where to look on a pad for emissions: equipment is ‘flagged’ for follow-up. A component-scale method is then employed to pinpoint the emission source and generate a ‘tag’. The system is designed to be implemented with the following workflow:

(1) Rapid screen: The PoMELO truck is immediately driven around the site upon arrival and a map of emission flags is produced. Quantification estimates can be performed if desired. This takes approximately 5 minutes but can be modified to extend longer if more accurate quantifications are required or the wind is light and variable.

(2) Component-scale survey: Individual units of equipment that were flagged during the rapid screen are surveyed with component-scale techniques such as M21 or OGI. Tags are placed following existing procedures familiar to regulators and operators. Repairs can be performed on the spot or queued for follow-up.

The advantage of the PoMELO workflow is accelerated application of component-scale surveys, reducing the total time on pad when portions are identified as not emitting by the initial equipment-scale screening. Additionally, the system produces an auditable record of data from the site.

The PoMELO system includes multiple survey modes with a variety of software components. The on-pad software used here is referred to as ‘Padmapper’. At the University of Calgary, we are developing different survey modes for the PoMELO hardware (or variations thereof). The results in this report do not apply to these other survey modes.

Figure 2: *The PoMELO system mounted on the roof of a standard field truck.*

Surveys at METEC were designed to benchmark the effectiveness of PoMELO in the rapid screen portion of the workflow only. We did not perform component-scale surveys for three reasons. First, component-scale surveys would add considerable time to each experiment, reducing the volume of data that could be collected to assess the screening approach. Second, component-scale surveys have an existing body of work examining their effectiveness (e.g., Ravikumar et al., 2018), and have been the subject of many new studies (Zimmerle, D., personal communication 13 Dec. 2019). The effectiveness of the rapid equipment-level screening approach is comparatively unknown. And third, emissions reductions equivalence with standard application of component-scale techniques (e.g., M21, OGI) is a leading metric used by regulators. If follow-up instructions from the PoMELO Padmapper successfully identify or 'flag' all emitting equipment and the distribution of leak rates in the experiment is representative, the work flow is very close to being equivalent in emissions reduction potential relative to a close-range only survey mode as the component-scale effectiveness is a common variable between both standard component-scale work flows, and the hybrid PoMELO approach.

Methods

METEC site overview

The METEC site was developed to assess the performance of technologies used to detect, localize, and quantify emissions on upstream oil and gas sites. The site consists of multiple mock oil and gas pads with real oil and gas equipment that has been modified to have hidden leak locations (Figure 3). There is no real oil and gas production on the site so the environment is completely controlled. Few people outside of METEC staff have knowledge of real leak locations.

Emissions are created from several centrally-located gas supply houses, from high pressure line natural gas. This gas was approximately 85% methane by mole fraction, determined from gas chromatagraph at Colorado State University and individually attributable to each gas supply system.

Emissions rates reported here are reported in 3 different units to ease reader understanding: (i) g/s CH_4, commonly used by scientists and the native units of the PoMELO system, (ii) scfh CH_4 (standard cubic feet per hour), commonly used by the US Natural gas industry (standard conditions: 70 °F, 1 atm.), and (iii) m^3/day CH_4, commonly used by Canadian regulators (standard conditions: 15 °C, 1 atm.). All rates are reported in terms of CH_4 only, any mixture of gas would have higher flow rates.

Each release point was precisely metered. The high pressure gas was initially cut down to lower pressures before feeding into a thermal mass flow meter. Downstream from this meter, the gas flow was run through a combination of up to 3 orifice plates to bring release rate(s) close to target(s). When multiple release points were fed from the same upstream flow meter, each release point was run in isolation to develop a calibrated flow rate estimate prior to opening all release points simultaneously for the experiment. These calibrations were done immediately prior to (but not during) each experiment to limit the effect of pressure or temperature changes on gas density and resultant choked flow through orifice plates.

Flow rates for each individual source were calculated as the flow fraction fed to each release point, as determined from pre-experiment calibrations multiplied by the settled flow rate metered for the full experiment. If one or more of the calibration flow rates did not settle to specification or was unable to be used, the calibrated flow rates for the known emissions points were used in place. Test 99 had an issue with settled flow rates, which required use of calibrated flow rates instead of the flow fraction method mentioned above. See Barchyn and Hugenholtz (2020) for full data and calculation process.

On 21 November 2019, calibrations were not run because ambient temperatures were below the required ambient operational conditions of the flow meters. On this day, replicates of previous experiments were run such that the exact same combinations of orifice plates and upstream pressures were used. Corrections were performed to account for the differences in temperature and pressure on choked flow through the orifice plates and compared against the measured aggregate flow rate. This results in slightly higher uncertainty on release rates, but aids in understanding PoMELO system performance at lower temperatures, which are common in Northern U.S. and Canada.

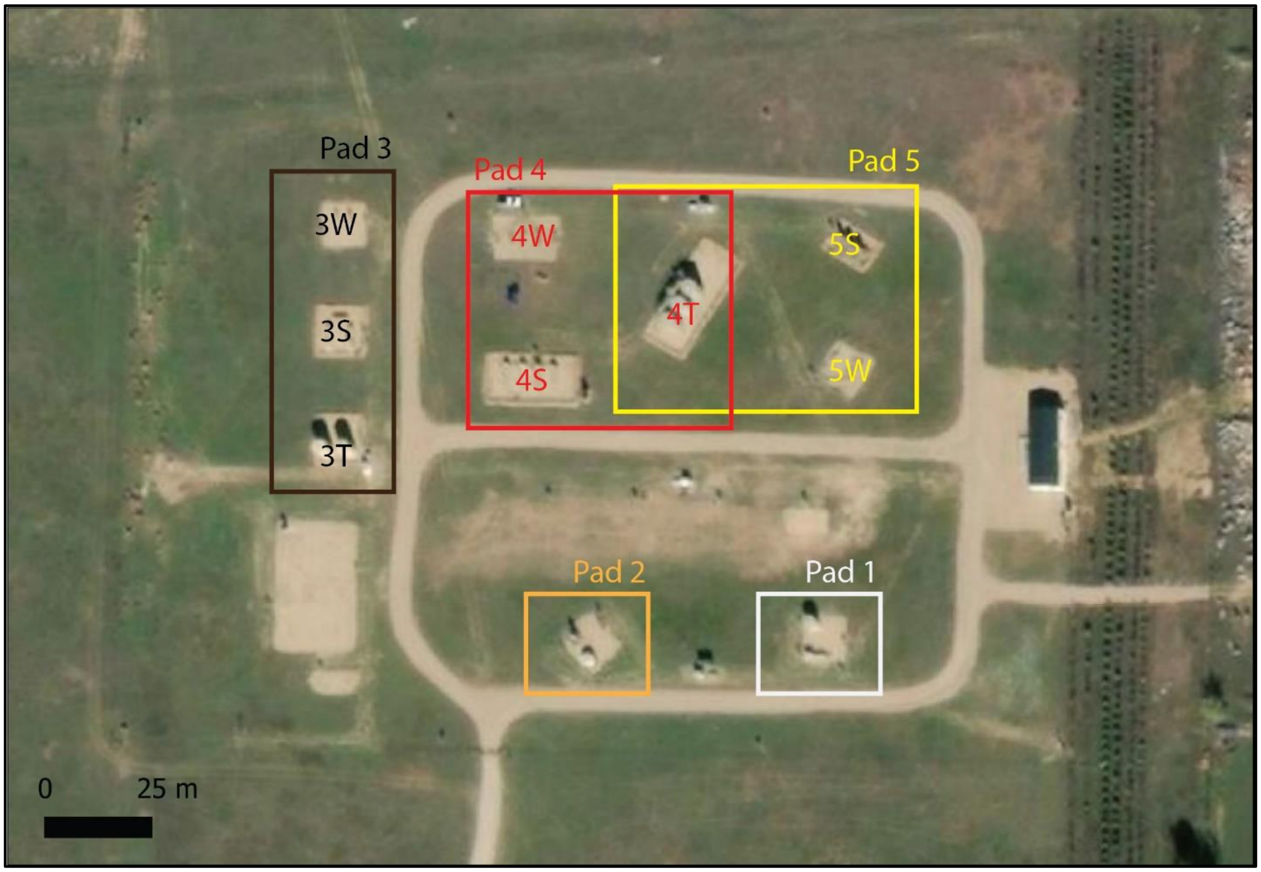

Figure 3: *Overview of different pads at METEC and the facility. Each pad contains multiple equipment units, typically grouped in like equipment. Pads 1 and 2 have one equipment group only.*

The following individual equipment units are present at the site and form the basis of our evaluations (Table 1). On each equipment unit there are many individual release points that a component scale survey would be able to survey.

Pad	Equipment Group	Equipment Unit
1	1 (wellhead, separator, tank)	1T-1 (tank)
		1S-1 (separator)
		1W-1 (wellhead)
2	2 (wellhead, separator, tank)	2T-1 (tank)
		2S-1 (separator)
		2W-1 (wellhead)
3	3W (wellheads)	3W-1 (wellhead)
		3W-2 (wellhead)
		3W-3 (wellhead)
	3S (separators)	3S-1 (separator)
		3S-2 (separator)
	3T (tanks)	3T-1 (tank)
		3T-2 (tank)
4	4W (wellheads)	4W-1 (wellhead)
		4W-2 (wellhead)
		4W-3 (wellhead)
		4W-4 (wellhead)
		4W-5 (wellhead)
	4S (separators)	4S-1 (separator)
		4S-2 (separator)
		4S-3 (separator)
		4S-4 (separator)
	4T (tanks)	4T-1 (tank)
		4T-2 (tank)
		4T-3 (tank)
5	5W (wellheads)	5W-1 (wellhead)
		5W-2 (wellhead)
		5W-3 (wellhead)
	5S (separators)	5S-1 (separator)
		5S-2 (separator)
		5S-3 (separator)
	4T (tanks)	4T-1 (tank)
		4T-2 (tank)
		4T-3 (tank)

Table 1: *Equipment hierarchy and listing. Each experiment was conducted on a given pad. Results are aggregated by equipment group and individual equipment units. Each equipment unit subsequently contains multiple emissions points that are not analyzed here. Note that Pad 5 includes the tank complex from Pad 4.*

The experiment pads differ in size and shape. Pads 1 and 2 are small and clustered, with only 1 equipment group each. The tanks present on Pads 1 and 2 are smaller (~100-200 barrel). Pads 3, 4, and 5 each contain 3 different equipment groups with much larger tanks (~300-400 barrel).

All pads have driving access around all sides, and around each equipment group. Ground level or overhead piping between equipment groups, which may hinder accessibility at real oil and gas facilities, did not exist at METEC during testing. This is important as PoMELO measures the methane concentrations in situ and must be downwind of any emitting equipment. This is enforced in the PoMELO software such that any infrastructure that is not upwind of the system cannot be reported as emitting or not emitting due to insufficient confidence.

Experiment protocol

To assess the PoMELO rapid-screen workflow, we used the following experimental protocol, similar to other pad-scale tests performed at METEC (e.g., Ravikumar et al., 2019). This protocol was designed to be blind to the PoMELO team to ensure any knowledge of the emissions configurations could not influence reporting.

An overview meeting was performed prior to any experiments. Both the PoMELO team and METEC team reviewed the experiment concept and goals. Some general target release rate ranges were discussed to ensure the experiment would yield useful results, but the PoMELO team was unaware of any specific pad configurations.

For each pad experiment, the METEC team created a synthetic emissions profile with 0-6 emissions sources operating at steady flowrates. Some pad configuration replicates were run. The test configurations were not known by the PoMELO team. During the pad setup and calibration phase, the PoMELO team was parked away from the test pad. Pads were tested in the following order 1, 3, 5, 2, 4, then repeated. This was to ensure the next pad was far away from the last pad and help ensure an equal number of tests were performed on each pad. In cases where the upcoming pad was downwind of the target pad and interference risk was minimal, calibrations could be run while PoMELO was surveying the target pad, enhancing testing efficiency. Upwind interference was actively monitored by the PoMELO team to ensure any unintended interference from non-target sources was minimal.

Upon the signal to begin the survey, the PoMELO team approached the pad, performed the survey, and used the output from the onboard software to report the METEC equipment pieces that were detected to be emitting, and should be flagged for follow-up, emulating real application. The elapsed time between when the Padmapper system started and stopped collecting data was used as a data collection elapsed time. Approximately 1-2 minutes was required after the system stopped recording data to fill out reporting datasheets. The total elapsed time for operations of the system can thus be taken as the data collection time, plus additional reporting time. Reporting time in practice will vary considerably depending on reporting systems.

The basis of reporting about both equipment units and equipment groups was the METEC equipment codes (see Table 1). Some pad configurations had multiple release points from a given piece of equipment. We analyzed data on both equipment unit and equipment group basis. Equipment unit comparisons represent a higher spatial resolution view of detection performance and measure the ability of the system to highlight which piece of equipment in a cluster is emitting. Equipment group comparisons provide a lower spatial resolution measurement of detection performance and measure the ability for the system to identify if an equipment cluster is emitting.

Testing was not paused or postponed for weather conditions. We provide further details on weather conditions during tests in the results section. All weather condition variables were measured onboard the

PoMELO vehicle system and are averages and standard deviations of the data collected over the data collection elapsed time. Wind speed and direction measurements were measured with the mobile anemometry algorithms onboard the PoMELO vehicle system.

Following reporting of results, the real pad configurations were provided to the PoMELO team whom subsequently performed all analysis. Consequently, the experiments were blind to the PoMELO team and provide a more reliable and defensible set of results (see Appendix A for details on our statement of blind test experiment process, and Appendix B for a statement of contributions). Full data are available in our downloadable supplementary data (Barchyn and Hugenholtz, 2020).

The PoMELO system was held constant throughout the experiments, no parameters were adjusted or otherwise modified. The system is designed to have no user-adjustable parameters such that a version of the hardware and software can be held constant and linked to a performance evaluation such as this one.

The specific version of hardware and software used in these experiments is identified internally with the code 'METEC_2019'. Screening systems are very sensitive to parameters; small parameter changes can create very large differences in results (Barchyn et al., 2019). The results here would not be valid if any consequential changes to the instrumentation or software were performed.

We do not include quantification data in this report as these data are not likely to be relevant for LDAR operations at present. We work forward with the workflow of following up on *all* detected anomalies. Triaging or other quantification-based efficiency enhancement schemes require significant external modeling to prove efficacy. Results are very dependent on poorly understood leak size distributions. Presently, we suggest the approach of following up on all detected anomalies is more defensible.

Analysis

Detections

Reported results include the METEC equipment codes that were detected to be emitting, with an equipment-specific or equipment-group specific emissions rate estimate.

	Emissions point present	**Emissions point absent**
Flagged (follow-up suggested)	(A) Inspection with close-range technique, with final emissions point probability of detection following existing experiments.	(B) Additional cost incurred for close-range inspection of equipment that is not emitting.
Not flagged (follow-up not suggested)	(C) Emissions source(s) was missed, resulting in a decrease in emissions reduction potential.	(D) Cost savings available as no close-range surveys are required.

Table 2: *Table of outcomes for PoMELO follow-up instructions for each piece of equipment or equipment group. The practical consequences of each outcome are discussed further in the text.*

Table 2 presents possible outcomes from implementation of the method where equipment is either flagged or not flagged for follow-up inspection when emission sources are either present or not present on the equipment in question. To further explain how the outcomes should be interpreted within the analysis, we briefly overview the practical ramifications of each outcome. In short, the most optimum results would be where all equipment either falls in outcome A or D. This demonstrates accurate emissions detection and localization as the system is accurate in separating equipment that is emitting from equipment that is not. The maximum cost savings occur as close-range surveys are as targeted as possible.

However, in practice there are inevitable localization and detection errors. Equipment falling in outcome B represents extra follow-up, which incurs extra costs as more close-range surveying is required than is necessary. Equipment in outcome C represents emissions points that are not identified during equipment-level screening and would reduce the total emissions reduction potential of the system for a given pad visit.

A full optimization of the different outcomes is beyond the scope of this report, but these outcomes are the guiding set of metrics for improving the system as there are direct linkages between each outcome and emissions reductions and cost. Optimizing the emissions reductions as a function of application cost is the overarching goal of LDAR technologies and we use this framing to guide our analysis and reporting of results herein.

Results and discussion

Experiment conditions

To provide context to the performance data, we report the experiment conditions. As a system dependent on the free atmosphere advecting methane from the sources to the sensors, of all atmospheric conditions, PoMELO is most sensitive to the wind conditions.

Experiments were run over 5 days: 18 Nov 2019 (Monday) (n = 30), 19 Nov 2019 (Tuesday) (n = 15), 20 Nov 2019 (Wednesday) (n = 20), 21 Nov 2019 (Thursday) (n = 20), and 22 Nov 2019 (Friday) (n = 20). Experiments were run the full day on Monday, and mornings Tuesday - Friday (other unrelated experiments were run in the afternoons).

Generally, earlier in the week, the weather was warm with light winds. Later in the week, the weather became snowy and cold (Figure 4). Winds were strongest on Wednesday and generally quite light through the rest of the week (Figure 5), but variable in direction (Figure 6). There was only trace precipitation over the week. We report environmental data on a per experiment (pad) basis.

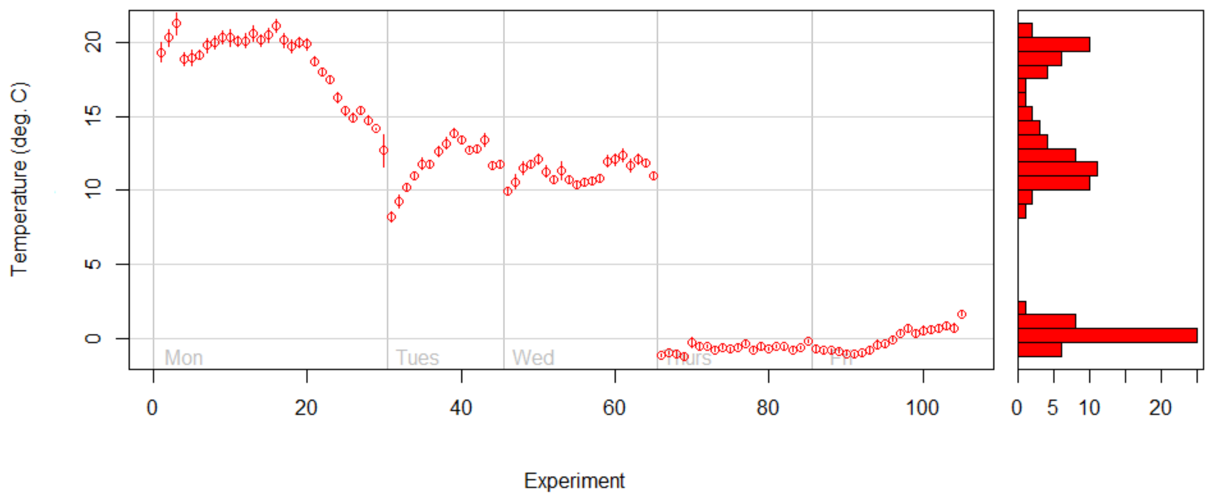

Figure 4 (above): *Mean experiment temperature ± 1 standard deviation. The right histogram shows the separation of warmer experiments (Monday - Wednesday), and colder experiments (Thursday - Friday).*

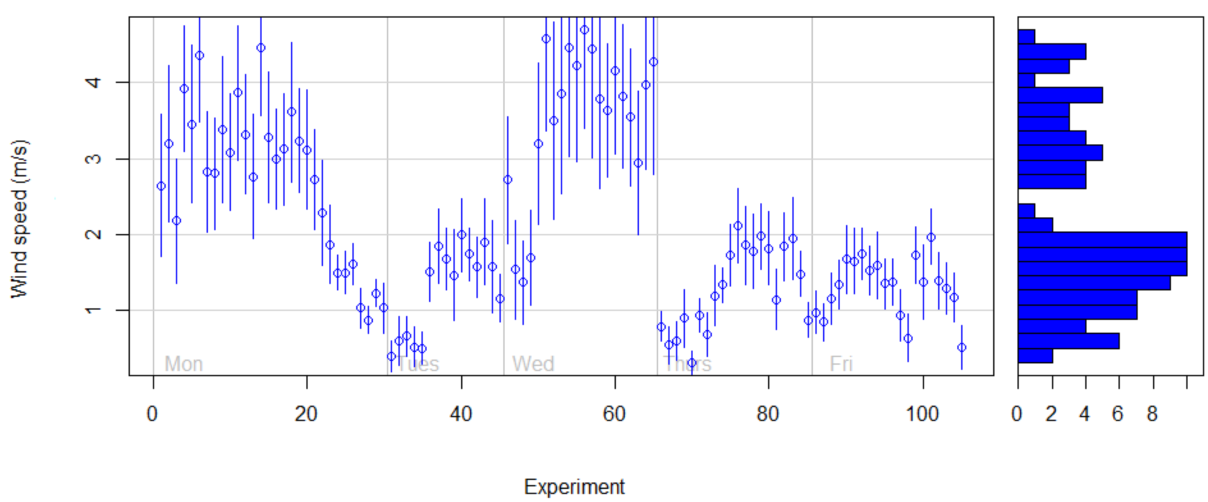

Figure 5 (above): *Mean experiment wind speed ± 1 standard deviation. The experiments with strongest wind speeds occurred on Wednesday, other days had lighter winds.*

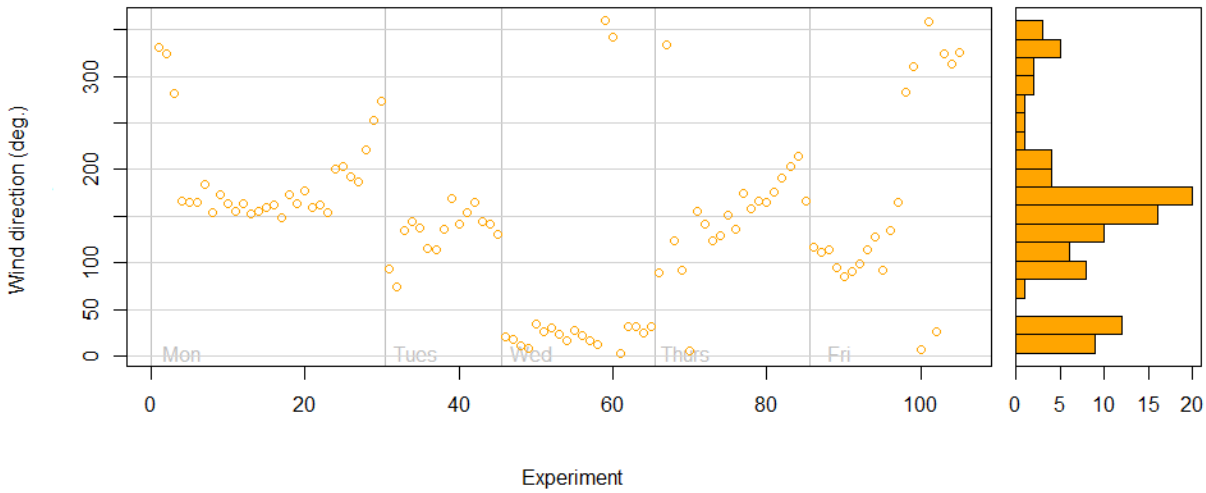

Figure 6 (above): *Mean experiment wind direction (azimuth wind was coming from). Most winds were from the S (~180°), but stronger winds on Wednesday were generally from the N. Winds were variable throughout most days.*

Fortunately, the test envelope was reasonably broad. However, we were unable to test the system in colder temperatures typical of Canadian winters (< -5 °C), or in stronger winds (> 5 m/s). This noted, we were able to obtain a broad sample of atmospheric conditions. Light and variable winds are common on upstream oil and gas pads in forested areas (where the on-pad winds are sheltered and affected by deeper boundary layers).

We did not quantify atmospheric turbulence characteristics. Although METEC itself is an open field, the surrounding surface roughness and topography is diverse. METEC is approximately 2 km E of the front range foothills of the Rocky Mountains. To the S and E, the city of Fort Collins presents considerable surface roughness with reasonably flat topography. To the N, a natural area represents a smooth surface roughness representative of many prairie environments.

The equipment onsite affected airflow, but in a similar manner to real oil and gas infrastructure. Of note, all equipment was open (similar to most infrastructure in the United States or other warm climates). Separators were not in small buildings, as is common in Canada or other colder climates. Tanks were clustered, bermed, and varied in size as is common on upstream oil and gas sites. Open equipment presents less airflow obstruction and results in less lee flow recirculation. Recirculating flow in the lee of obstructions is a major analytics challenge as the wind direction is variable and simply drawing a line upwind from a plume tends to be less accurate. As such, the presence of real equipment, and the effects of this equipment on airflow is an essential advantage of METEC. Similar experiments with release stacks and no local interference of airflow are not as representative.

PoMELO survey characteristics

The PoMELO system was driven in a normal survey manner for each experiment to emulate normal operations, with > 3 laps of each equipment group (Figure 7 shows survey times). We drove all pads at a normal driving pace for upstream pads (Figure 8).

The survey time and driving speed is internally enforced in the PoMELO software through a running measurement of confidence. Additionally, safety policies of most upstream producers mandate a maximum driving speed. The minimum survey time was 40 s, the maximum survey time was 11.0 minutes. All experiments significantly surpassed the minimum enforced confidence.

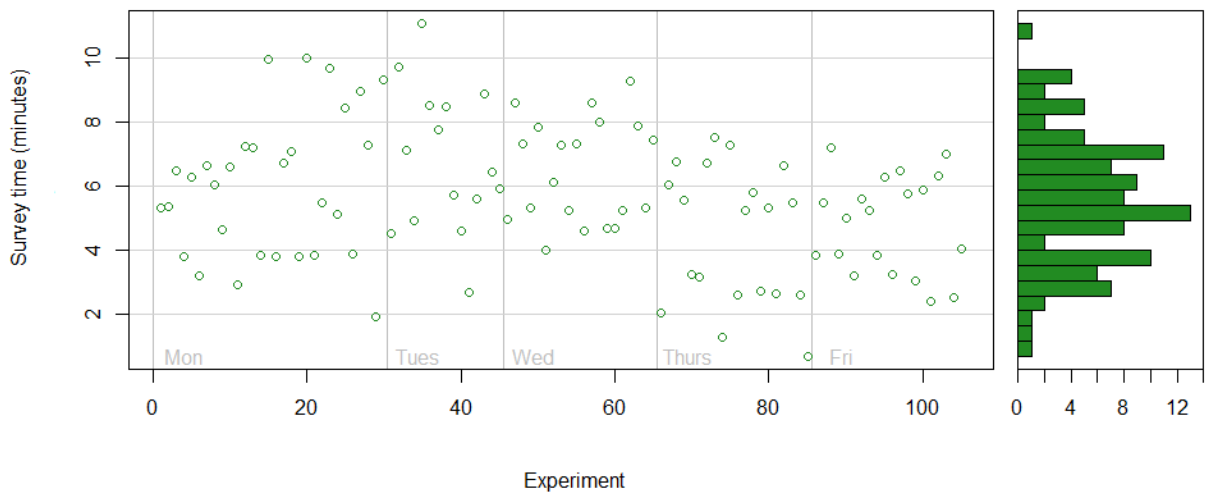

Figure 7 (above): *Survey time for each experiment. Note that survey time is recorded by the system and does not include time for reporting which is completed after turning off the system and is a fixed time interval. Note that pads are not equal sized and longer survey times do generally correspond to the larger pads (Pad 3, 4, 5).*

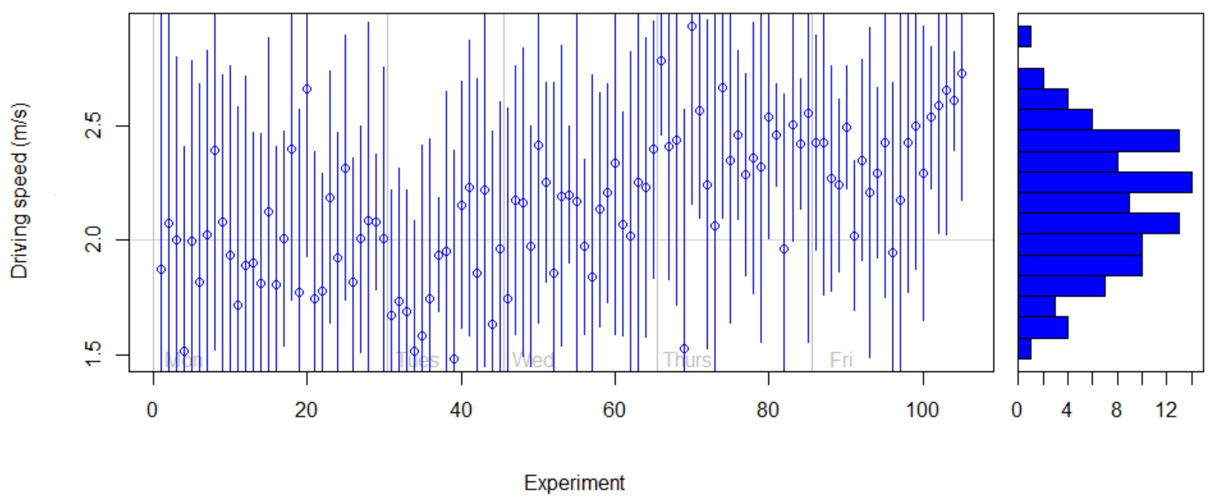

Figure 8 (above): *Driving speed for each experiment ± 1 standard deviation. For reference 1 m/s = 3.6 km/h = 2.2 mph.*

Blind pad characteristics

Prior to reporting measured results, we context the blind pad characteristics that the PoMELO system was subjected to (Figures 9-13). These characteristics were designed to emulate similar experiments (Ravikumar et al., 2019), but also were designed to explore the limitations of the PoMELO system. We did not conceptually or practically separate vents from leaks – this is done at the follow-up stage with close range survey methods and is outside of the scope of these experiments.

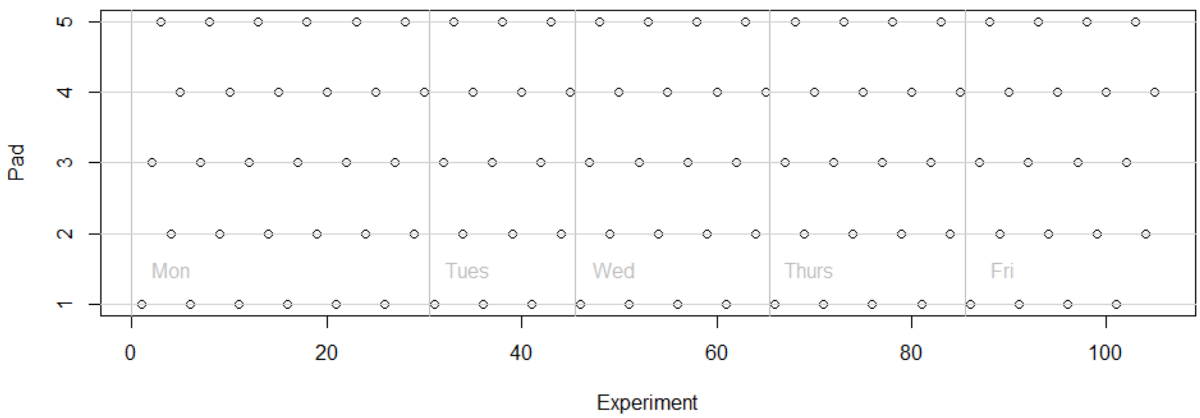

Figure 9 (above): *Pad distribution for each experiment. Pads were run in the following order 1, 3, 5, 2, 4 to limit any effects of individual pads and facilitate maximum spatial separation between the last pad completed and next pad to be surveyed. This order was maintained throughout all experiments.*

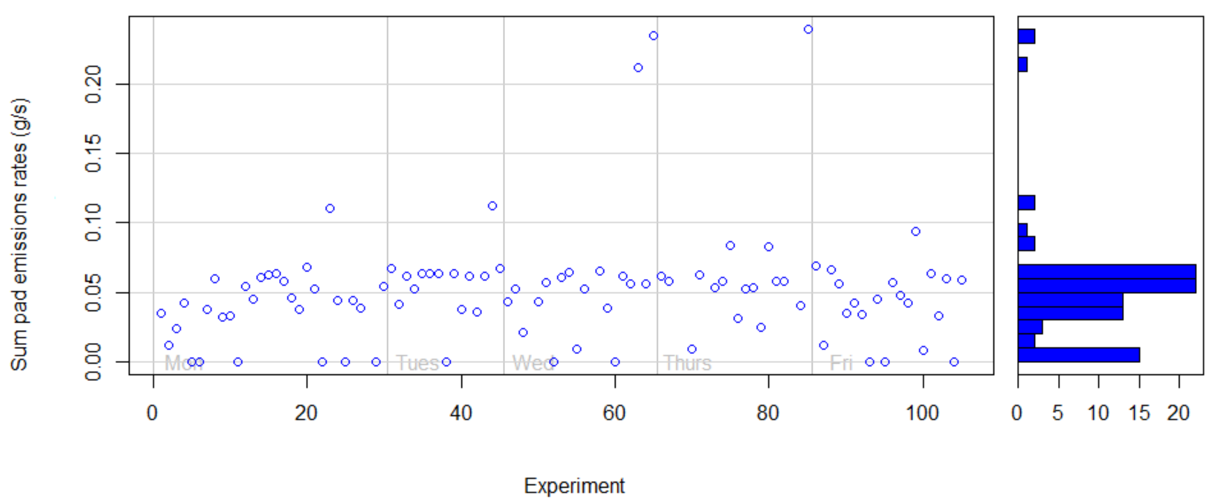

Figure 10 (above): *Sum of emissions rates for each experiment pad in g/s CH_4. Refer to Figures 11 and 12 for different units.*

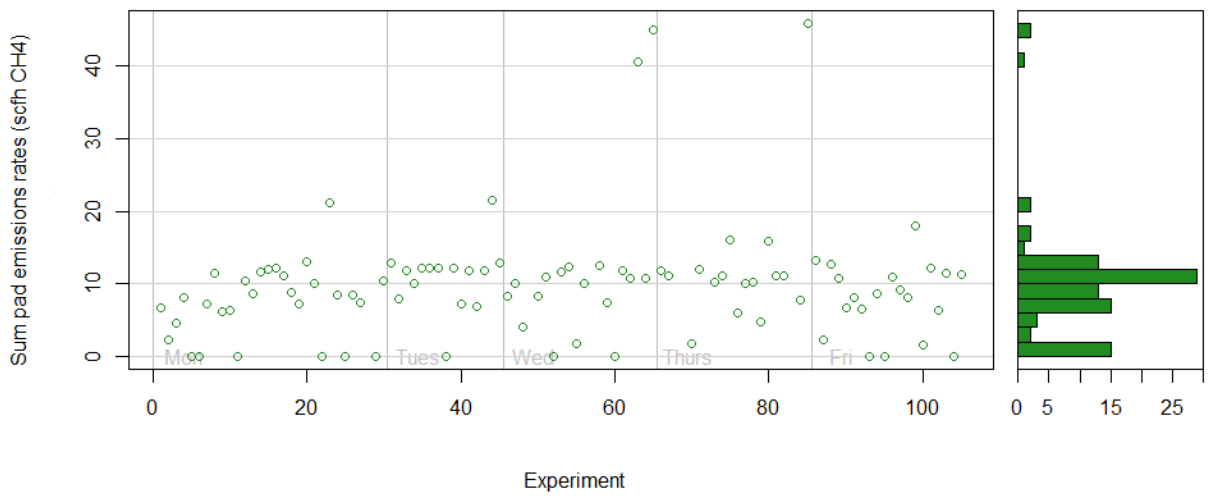

Figure 11 (above): *Sum of emissions rates for each experiment pad in standard cubic feet per hour (scfh) CH_4. Included here to help readers not familiar with other units.*

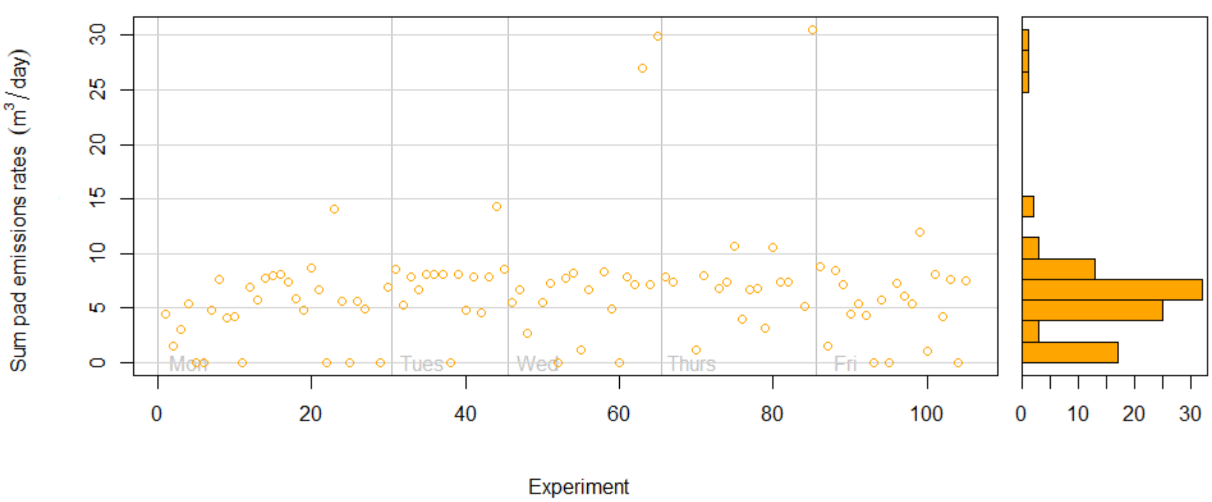

Figure 12 (above): *Sum of emissions rates for each experiment pad in m^3/day CH_4. Included here to help readers not familiar with other units.*

The total pad emissions rates are generally similar to those of Ravikumar et al. (2019). These rates are much lower than empirically measured total pad emissions. For example, Zavala-Araiza et al. (2018) reported average total pad emissions of 1.52 g/s in Alberta, Canada, measured empirically from off-site. We used much lower rates as the PoMELO system is quite sensitive to small emissions rates and we wished to benchmark the skill of the system at these lower rates.

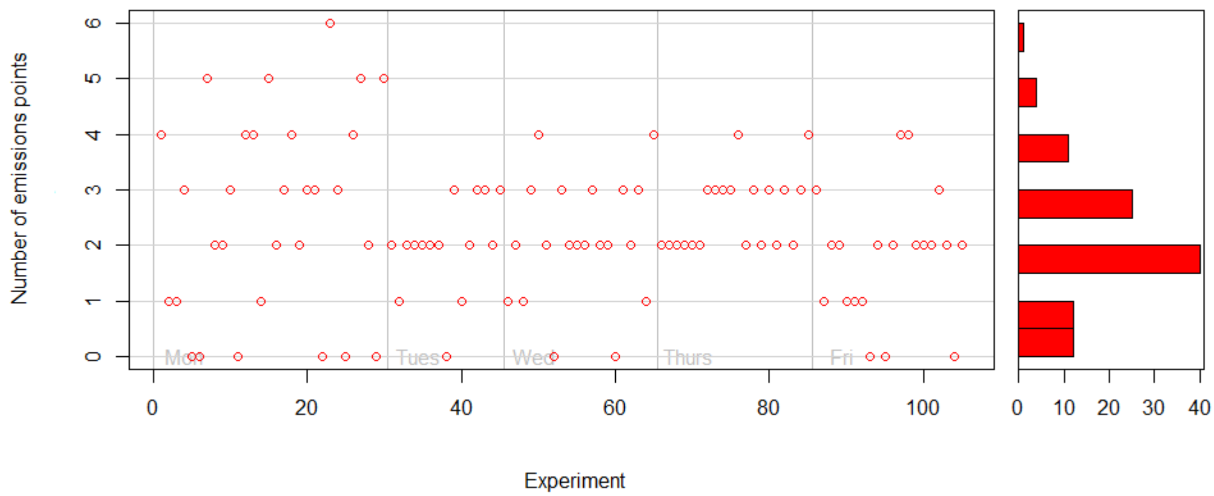

Figure 13: *The number of emissions points on each pad was varied between 0 and 6, with most pads containing 2 emissions points.*

We report two main analysis streams, equipment and equipment group. Refer to Table 1 to link pads, equipment, and equipment groups. To better understand the proportion of available equipment (or equipment groups) that were emitting, we examined the ratio of $n_{emitting} / n_{available}$. This shows the fraction of available equipment that were emitting (Figures 14), or fraction of available equipment groups that were emitting (Figure 15). When this ratio is 1.0, all possible equipment (or equipment groups) were emitting, when this ratio is 0.0, nothing was emitting.

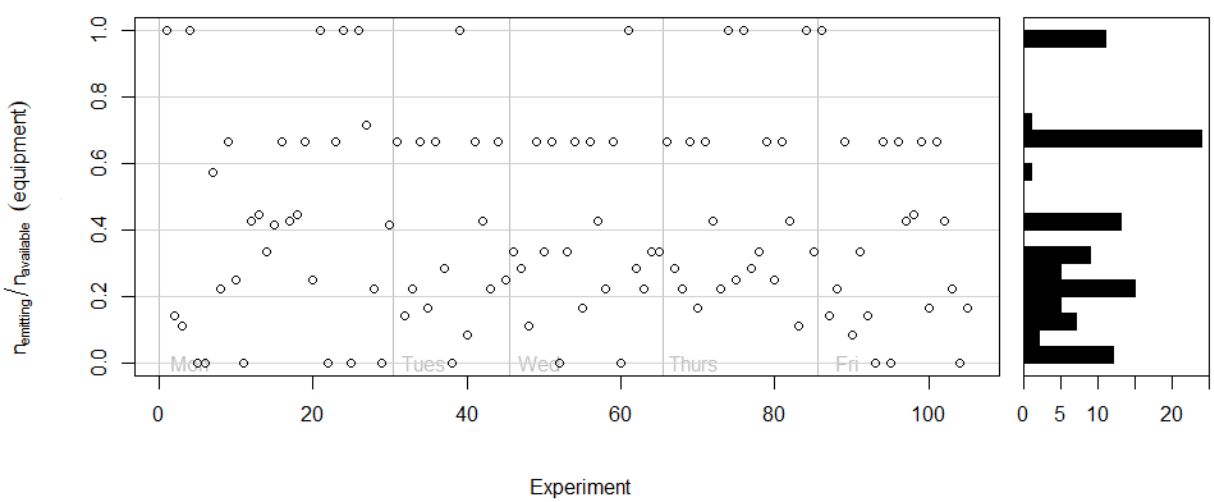

Figure 14 (above): *The proportion of equipment within an experiment with at least one emission point. It was infrequent that all equipment was emitting. Most experiments had a small fraction of emitting equipment.*

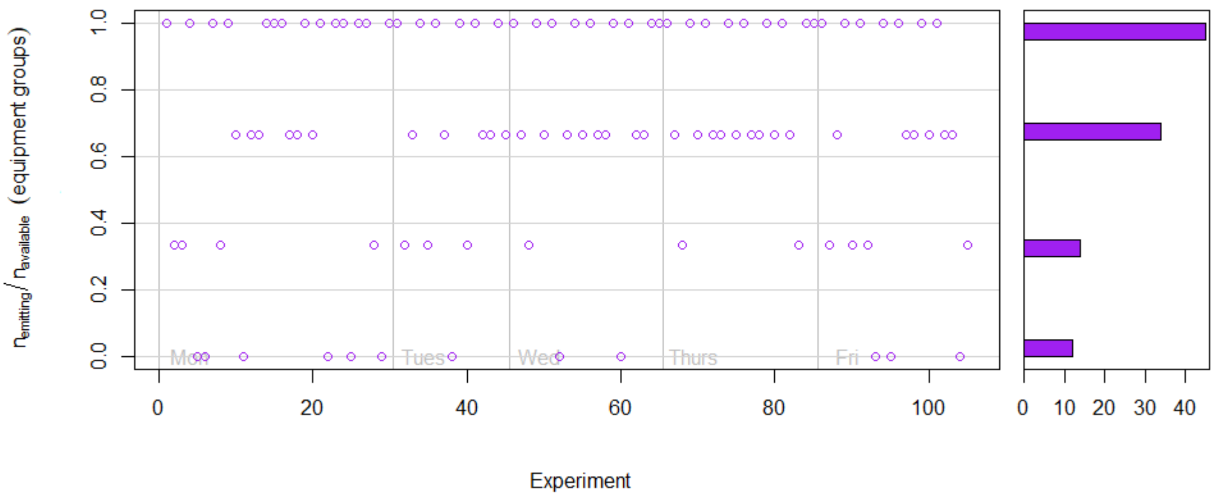

Figure 15 (above): *The proportion of equipment groups within an experiment that were emitting. Note that Pads 1 and 2 only had 1 equipment group, so this plot shows a binary result: 0.0, not emitting, or 1.0, emitting. Pads 3, 4, and 5 and 3 equipment groups each so they either plot at 0.0, 0.33, 0.67, or 1.0.*

It is unclear if the emissions configurations of pads used in this experiment are precise representations of typical upstream oil and gas pads, although the leak locations on METEC equipment are broadly representative of the typical places on oil and gas equipment that leak and vent. The emissions configurations of upstream pads are both variable and poorly understood. Some jurisdictions enforce venting regulations or stringent LDAR requirements that could result in upstream pads with fewer release points than jurisdictions with little regulation. For example, venting can dominate emissions if production is targeting liquids and there is little opportunity to tie in waste gas. Consequently, there is no typical upstream pad configuration, and little available data to support precision pad configuration. Further, all emission sources were operated at a steady emission rate during experiments, and intermittent or unsteady venting was not simulated in the pad configurations. The pad configurations tested here provide a reasonable compromise between replicating upstream production and providing targeted data to probe the capabilities of the PoMELO system.

Although the experiment pad aggregation level emulates real oil and gas pads, both the equipment (Figure 16) and equipment group data (Figure 17) can be analyzed outside the context of pad experiments, and we do this throughout this report where required.

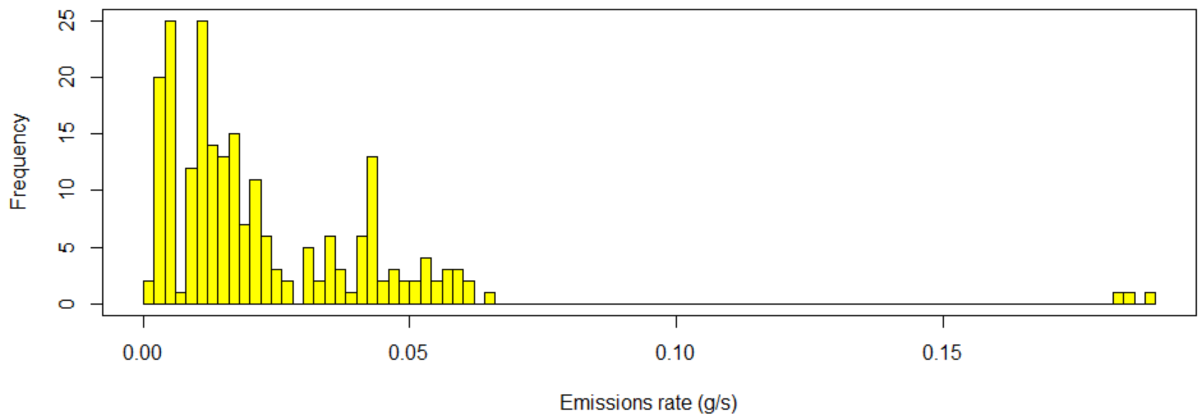

Figure 16 (above): *The distribution of equipment unit aggregated emissions rates, aggregated among all pad experiments.*

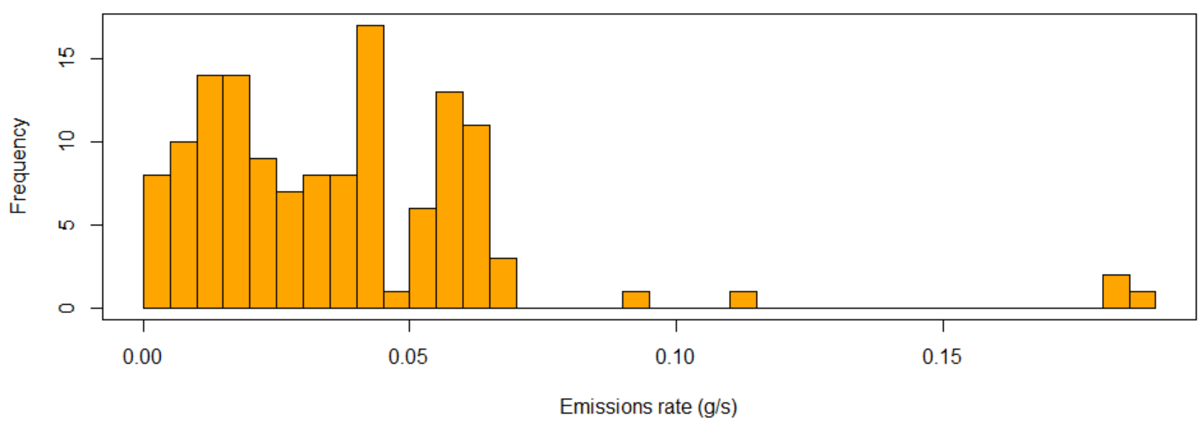

Figure 17 (above): *The distribution of equipment group aggregated emissions rates, aggregated among all pad experiments.*

PoMELO detection metrics

PoMELO provides detections to trigger close-range follow-up. The metrics can be framed in terms of the system's ability to correctly detect emitting equipment, or equipment groups for follow-up (Figures 18 and 19). Correct detections closely relate to the emissions reduction potential of the system. The highest emissions reduction potential would occur with 100% correct flags.

Unnecessary detections represent extra follow-up, with an efficiency and cost penalty. Note that this efficiency penalty manifests as extra close-range follow-up. We metric the proportion of unnecessary detections to better inform the cost reductions potential of the system and benchmark the ability for the system to 'not detect' emissions points. Both the ability to correctly detect emissions points, and reduce unnecessary detections are central metrics to demonstrate the performance of the PoMELO system in its design deployment.

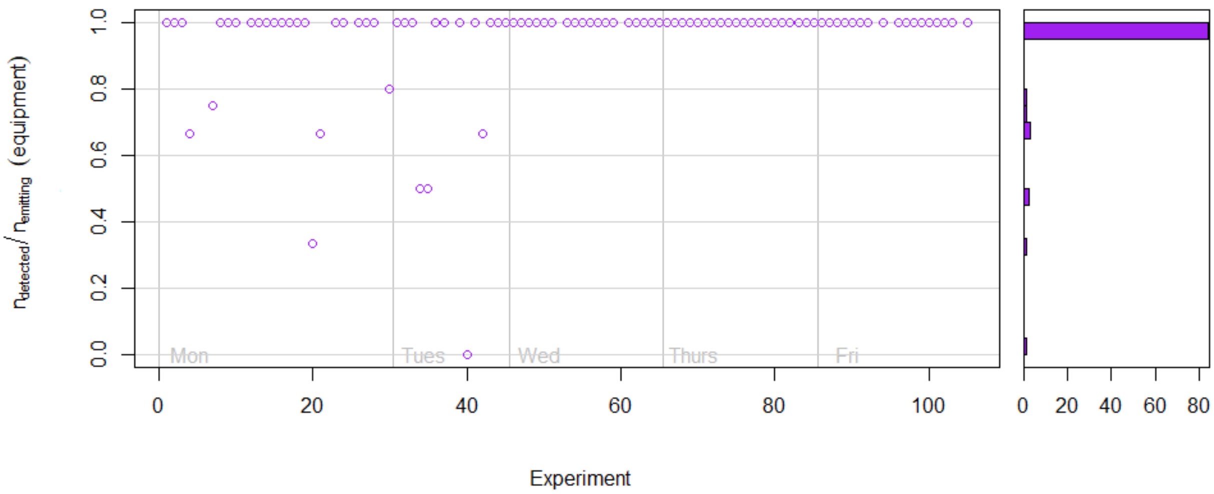

Figure 18 (above): *The proportion of correctly detected equipment units by experiment as the ratio of $n_{detected}/n_{emitting}$. In total, 90.3% of experiments had all emitting equipment units correctly flagged for follow-up (n = 93). Experiments with no emitting equipment units are not included.*

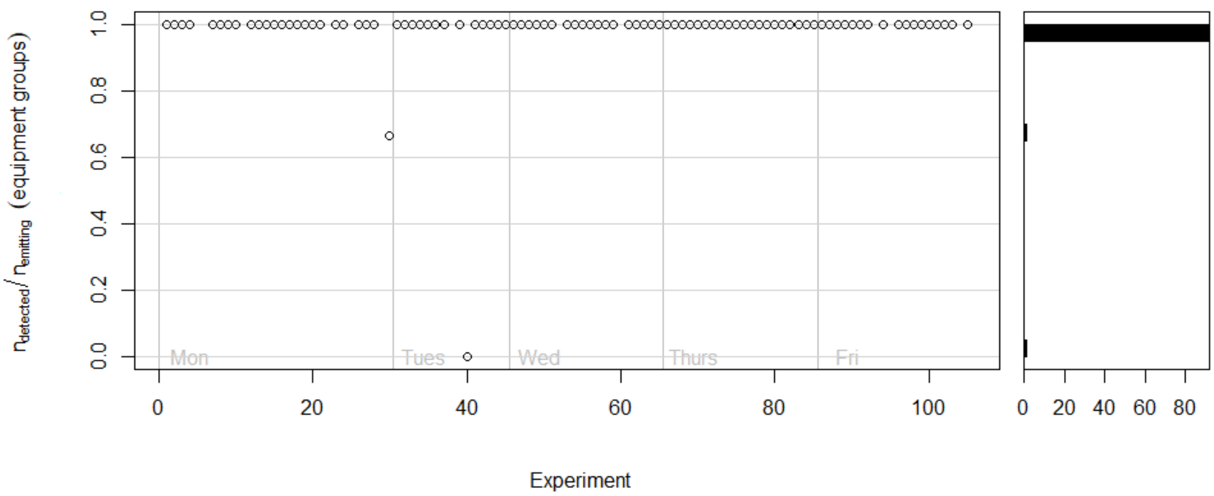

Figure 19 (above): *The proportion of correctly detected equipment groups by experiment as a ratio of $n_{detected}/n_{emitting}$. In total, 97.8% of experiments had all emitting equipment groups correctly flagged for follow-up (n = 93). Experiments with no emitting equipment groups are not included.*

The PoMELO system successfully detected most of the emitting equipment on sites. Only one experiment occurred where the PoMELO system was unable to detect any emissions points when there was an emissions point on the pad. Figures 20 and 21 show the fraction of clean equipment and equipment groups that would be subject to close-range surveys.

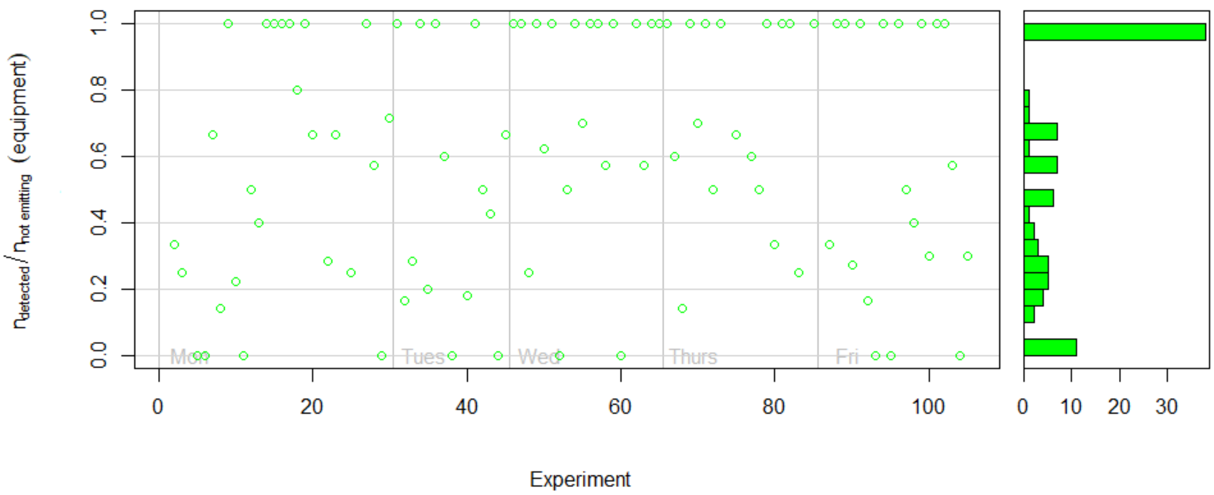

Figure 20 (above): *The proportion of extra equipment scale flags for each experiment as a ratio of $n_{detected}$ / $n_{not\ emitting}$. A significant number of equipment was flagged for close-range follow-up that was not emitting.*

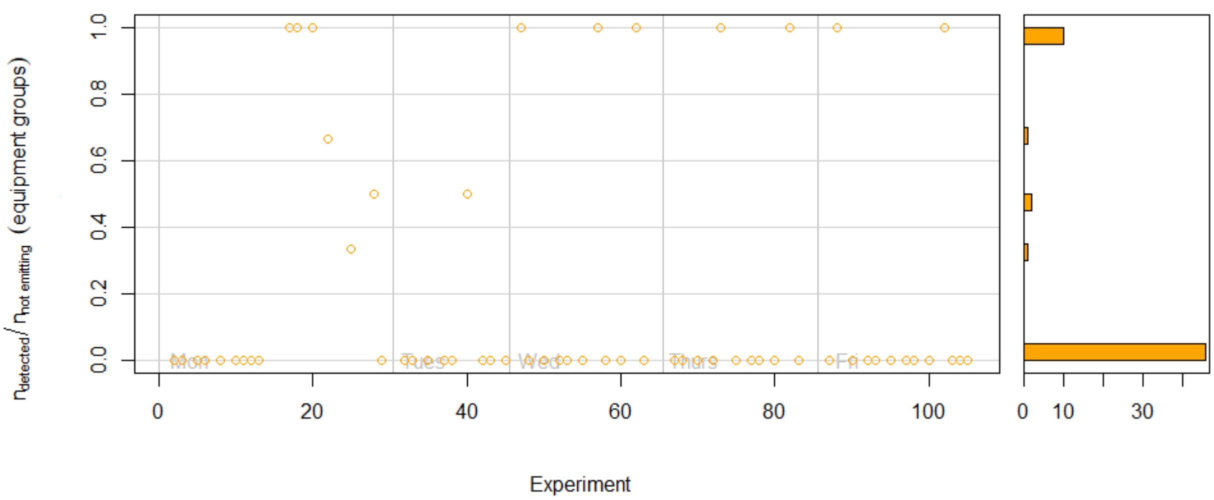

Figure 21 (above): *The proportion of extra equipment group detections for each experiment as a ratio of $n_{detected}$ / $n_{not\ emitting}$. These results suggest localization accuracy and ability to identify infrastructure that is not emitting is much better at the equipment group scale than at the equipment scale (see Figure 20).*

To better understand the capacity for PoMELO to identify emitting and non-emitting equipment correctly, we can pull the equipment and equipment group data out of each pad experiment and analyze it as a single population (Tables 3 and 4).

	Emissions present	**Emissions absent**	**SUM**
Follow-up suggested (detected)	216	230	446
Follow-up not suggested (not detected)	10	258	268
SUM	226	488	714

Table 3 (above): *Detection metrics analyzed at the equipment unit scale.*

	Emissions present	Emissions absent	SUM
Follow-up suggested (detected)	139	15	154
Follow-up not suggested (not detected)	2	75	77
SUM	141	90	231

Table 4 (above): *Detection metrics analyzed at the equipment group scale.*

In general, the PoMELO system was sensitive and detected emissions reliably. This suggests the emissions reduction potential of the system is robust. However, the system did flag many individual pieces of equipment that were not emitting. This is an efficiency penalty as this equipment would be surveyed with close-range methods and no emissions would be found.

The results were not consistent between equipment and equipment groups, suggesting these higher follow-up rates were caused by the spatial scale of detections. In other words, it is likely that if there was an emissions point in an equipment group, each equipment unit in the group was labeled as emitting as it was not clear which individual unit was emitting.

This bias towards more follow-up instead of less follow-up is to some extent deliberate and was likely by design as, although extra follow-up represents an efficiency penalty, it is likely in most situations that the expense associated with extra follow-up is minimal compared to missing emission points, particularly if other portions of the pad are emitting as well.

Additionally, these results are likely modulated by the specific spatial scale of equipment and equipment groups at METEC (Figure 1). Equipment is grouped in small square configurations and access is good around each equipment. On real oil and gas pads, equipment is often grouped similar to METEC, but in other cases, equipment is more dispersed. This spatial clustering of equipment would modulate the efficiency of the PoMELO system in real practice. If all equipment is dispersed, the efficiency of the system would likely be enhanced as it is easy to isolate emitting equipment. If all equipment is clustered, the efficiency of the system is reduced as it is difficult to determine which pieces in a cluster are emitting and which are not. The wind direction relative to the position of the equipment also affects results. The best situation is where wind direction is perpendicular to clustering.

The ability to analyze the effect of detection skill of the PoMELO system on an equipment basis is an important characteristic of METEC. First, this emphasizes that pure timeseries detection (e.g., Barchyn et al., 2019) is insufficient in many cases to metric the skill of systems as it is important to be able to not only identify that there is an emissions source present, but also point to where it is likely to be occurring. From an algorithm perspective, this is likely more difficult in practice than finding anomalies. Second, the ambiguity in localization is likely closely related to the airflow characteristics found in the lee of large oil and gas equipment (in particular tanks), if similar experiments were conducted with ground release points that don't affect the nearfield airflow, localization results would likely be more accurate due to the lack of lee turbulence. These effects suggest the METEC approach is very valuable if results are designed to inform real deployments.

Minimum detection limits

A classical method of evaluating the performance of LDAR technology and methods is modeling the minimum detection limits (MDL) as a function of various controlling variables. The most common controlling variable is the emissions rate where it is expected that a curve would range from 0.0 to 1.0 over a range of emissions rates from too low to be detected, to those which are almost always detected as they are so large (e.g., Barchyn et al., 2019; Ravikumar et al., 2019). Experiments designed to comprehensively evaluate an MDL curve should span the range of emissions rates from undetectable to consistently detectable.

We analyze the MDL at both the equipment and equipment group scale, outside of the context of experiment pads. We bin the data into equal-sized bins of width 0.01 g/s (≈ 1.91 scfh CH_4, ≈ 1.27 m^3/day). This provides a reasonable sample for each bin to better understand detection probability (Figures 22-24).

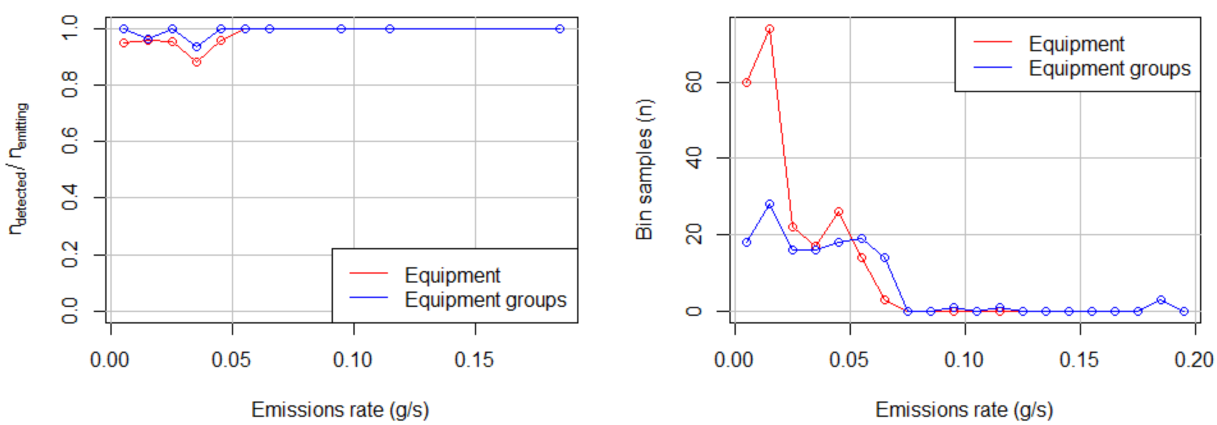

Figure 22 (above): *Detection frequency as a function of equipment emissions rate, with matching bin sample sizes. Points are plotted at the bin center along the x-axis (Table 5).*

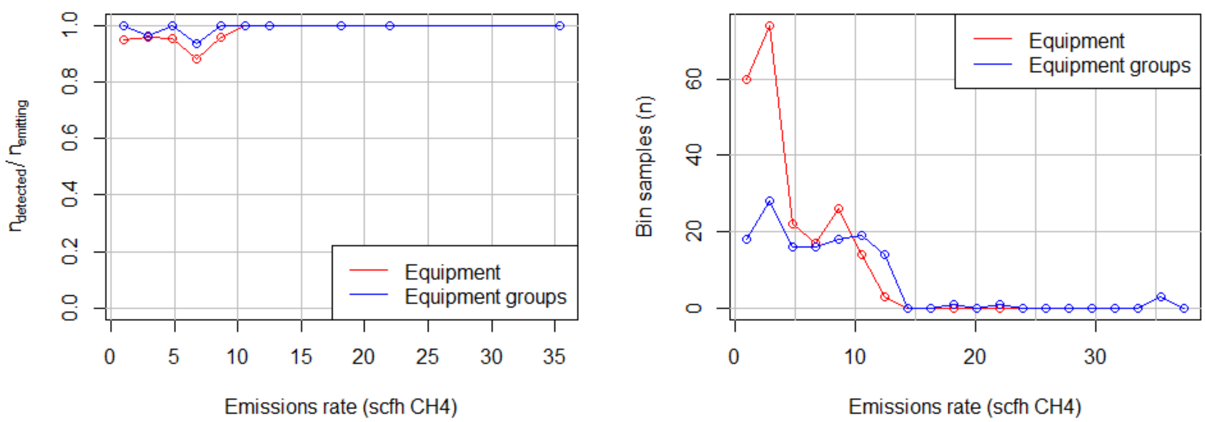

Figure 23 (above): *Detection frequency as a function of equipment emissions rate, with matching bin sample sizes. Points are plotted at the bin center along x axis (Table 5). Identical to Figure 20, but with units of scfh CH_4.*

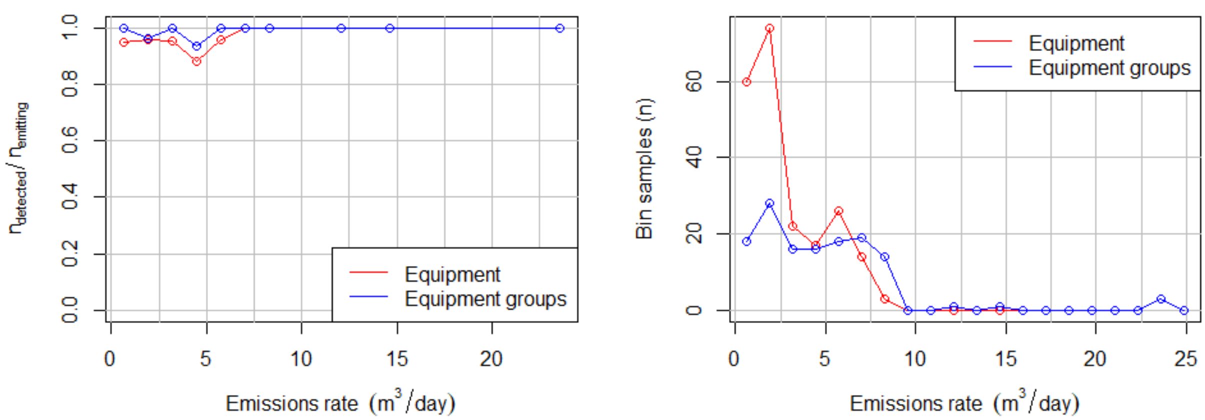

Figure 24 (above): *Detection frequency as a function of equipment emissions rate, with matching bin sample sizes. Points are plotted at the bin center along x axis (Table 5). Identical to Figure 20, but with units of m^3/day CH_4.*

Bins (g/s)	Number of equipment	$n_{detected}$ / $n_{emitting}$ (equipment)	Number of equipment groups	$n_{detected}$/ $n_{emitting}$ (equipment groups)
0.00 - 0.01	60	0.950	18	1.000
0.01 - 0.02	74	0.959	28	0.964
0.02 - 0.03	22	0.955	16	1.000
0.03 - 0.04	17	0.882	16	0.938
0.04 - 0.05	26	0.962	18	1.000
0.05 - 0.06	14	1.000	19	1.000
0.06 - 0.07	3	1.000	14	1.000
0.07 - 0.08	0	-	0	-
0.08 - 0.09	0	-	0	-
0.09 - 0.10	0	-	1	1.000
0.10 - 0.11	0	-	0	-
0.11 - 0.12	0	-	1	1.000
0.12 - 0.13	0	-	0	-
0.13 - 0.14	0	-	0	-
0.14 - 0.15	0	-	0	-
0.15 - 0.16	0	-	0	-
0.16 - 0.17	0	-	0	-
0.17 - 0.18	0	-	0	-
0.18 - 0.19	3	1.000	3	1.000
0.19 - 0.20	0	-	0	-

Table 5: *Detection statistics (plotted in Figures 22-24).*

The number of undetected equipment and undetected equipment groups was small (n = 10 and n = 2, respectively). These small sample sizes limit comprehensive understanding of causal factors. There is some bias towards missed detections with smaller emissions rates, but there are also more data at these smaller release rates suggesting infrequent or random issues not correlated to emissions rate could also preferentially appear in the data. These issues could include reporting errors or issues with upwind pads and overlapping plumes. Future analyses will explore these missed emissions to understand if the system can be improved or if a certain error rate is irreducible.

It is likely that to better understand the detection limit curves, much lower flow rates should be tested. This noted, the range of flow rates required may be below rates that are possible to meter at METEC. Evaluating a full detection limit curve was a secondary goal of this study but could form the focus of follow-up work.

The minimum emissions rate tested here was below 98.51% of emissions rates found by Clearstone Engineering and Carleton University in a comprehensive 2018 study of upstream emissions in Alberta, Canada (n = 402). Expressed in terms of sum emissions rate, 100.00% (99.99998% before rounding) of emissions identified by Clearstone Engineering and Carleton University were greater than the minimum tested here. Note these field emissions rates were measured in terms of total gaseous hydrocarbons and modulated with the measurement methods used (see Clearstone Engineering and Carleton University, 2018). Nevertheless, these data suggest the minimum rates tested here, although insufficiently low to fully map the MDL curve of PoMELO, were very low relative to most leaks identified and targeted in upstream LDAR programs. Additionally, the MDL curves of OGI cameras (the most common close-range follow-up method) sharply drop off before the minimum tested here (Ravikumar et al., 2018), indicating PoMELO is likely able to flag emissions rates that would be difficult to find with an OGI camera.

We can take a deeper look at the undetected equipment and equipment groups (Tables 6 and 7), but the lack of sample size limits our ability to make defensible inferences about the variables that are controlling missed detections.

Test	Name	Temperature (°C) (percentile)	Wind speed (m/s) (percentile)	Mean wind direction (deg.)	Equipment emissions rate (g/s) ± 1 std. dev.	Equipment emissions rate (scfh CH4) ± 1 std. dev.	Equipment emissions rate (m^3/day) ± 1 std. dev.
4	2S-1	18.87 (81%)	3.92 (92%)	165.56	$5.24 \times 10^{-3} \pm 4.4 \times 10^{-5}$	$1.00 \pm 8.4 \times 10^{-3}$	$6.68 \times 10^{-1} \pm 5.6 \times 10^{-3}$
7	3S-1	19.80 (86%)	2.82 (71%)	184.08	$2.19 \times 10^{-2} \pm 8.6 \times 10^{-4}$	$4.20 \pm 1.6 \times 10^{-1}$	$2.79 \pm 1.1 \times 10^{-1}$
20	4S-3	19.88 (87%)	3.12 (75%)	176.72	$1.21 \times 10^{-2} \pm 9.5 \times 10^{-5}$	$2.32 \pm 1.8 \times 10^{-2}$	$1.54 \pm 1.2 \times 10^{-2}$
20	4T-3	19.88 (87%)	3.12 (75%)	176.72	4.14 x 10-2 ± 3.3×10^{-4}	$7.93 \pm 6.3 \times 10^{-2}$	$5.28 \pm 4.2 \times 10^{-2}$
21	1T-1	18.72 (80%)	2.73 (66%)	160.04	3.72 x 10-2 ± 7.2×10^{-4}	$7.12 \pm 1.4 \times 10^{-1}$	$4.74 \pm 9.2 \times 10^{-2}$
30	4T-2	12.70 (65%)	1.03 (23%)	273.65	1.65 x 10-2 ± 3.1×10^{-4}	$3.16 \pm 5.9 \times 10^{-2}$	$2.10 \pm 3.9 \times 10^{-2}$
34	2W-1	11.00 (48%)	0.51 (4%)	143.85	5.77 x 10-3 ± 1.0×10^{-4}	$1.10 \pm 2.0 \times 10^{-2}$	7.35 x 10-1 ± 1.3×10^{-2}
35	4W-4	11.78 (56%)	0.50 (3%)	138.02	1.52 x 10-2 ± 2.6×10^{-4}	$2.91 \pm 5.0 \times 10^{-2}$	$1.94 \pm 3.3 \times 10^{-2}$
40	4S-4	13.40 (68%)	1.99 (62%)	141.88	3.76 x 10-2 ± 9.8×10^{-5}	$7.20 \pm 1.9 \times 10^{-2}$	$4.79 \pm 1.2 \times 10^{-2}$
42	3W-3	12.82 (67%)	1.57 (41%)	164.90	4.65 x 10-3 ± 3.0×10^{-4}	$8.91 \times 10^{-1} \pm 5.8 \times 10^{-2}$	$5.93 \times 10^{-1} \pm 3.9 \times 10^{-2}$

Table 6 (above): *Equipment that was not flagged but was emitting. Percentiles are calculated as the rank within the equipment surveyed.*

Test	Name	Temperature (°C) (percentile)	Wind speed (m/s) (percentile)	Mean wind direction (deg.)	Group emissions rate (g/s) ± 1 std. dev.	Group emissions rate (scfh CH4) ± 1 std. dev.	Group emissions rate (m^3/day) ± 1 std. dev.
30	4T	12.70 (69%)	1.03 (23%)	273.65	$1.65 \times 10^{-2} \pm 3.1 \times 10^{-4}$	$3.16 \pm 5.9 \times 10^{-2}$	$2.10 \pm 3.9 \times 10^{-2}$
40	4S	13.40 (72%)	1.99 (62%)	141.88	$3.76 \times 10^{-2} \pm 9.8 \times 10^{-5}$	$7.20 \pm 1.9 \times 10^{-2}$	$4.79 \pm 1.2 \times 10^{-2}$

Table 7 (above): *Equipment groups that were not flagged but were emitting. Percentiles are calculated as the rank within the equipment surveyed.*

The missed emissions points (Tables 6 and 7) do not show a clear alignment with certain conditions. Most issues occurred on Pad 4, but Pad 4 is not tangibly different than any other pad at METEC.

At the equipment scale, the total sum emissions rates flagged by PoMELO was 4.75 g/s, with 0.198 g/s missed (96.0 % of sum emissions rates flagged). At the equipment group scale the total sum emissions rates flagged by PoMELO was 4.90 g/s, with 0.0541 g/s missed (98.9 % of sum emissions rates flagged).

Broadly the PoMELO system shows very sensitive detection capabilities. Although the results from Ravikumar et al. (2019) cannot be quantitatively compared to these as the test protocols differed, the range of release rates and detection efficacy of PoMELO suggests that it has similar detection capabilities to competing systems.

References

Alvarez, R.A., Zavala-Araiza, D., Lyon, D.R., Allen, D.T., Barkley, Z.R., Brandt, A.R., Davis, K.J., Herndon, S.C., Jacob, D.J., Karion, A., Kort, E.A., Lamb, B.K., Lauvaux, T., Maasakkers, J.D., Marchese, A.J., Omara, M., Pacala, S.W., Peischl, J., Robinson, A.L., Shepson, P.B., Sweeney, C., Townsend-Small, A., Wofsy, S.C., Hamburg, S.P., 2018. Assessment of methane emissions from the U.S. oil and gas supply chain. *Science* 361, 186-188. DOI: 10.1126/science.aar7204.

Barchyn, T.E., Hugenholtz, C.H., 2020. PoMELO METEC 2019 Data. Supporting data. DOI: https://doi.org/10.7910/DVN/7ZZKC7.

Barchyn, T.E., Hugenholtz, C.H., Fox, T.A., 2019. Plume detection modeling of a drone-based natural gas leak detection system. *Elementa: Science of the Anthropocene* 7, 41. DOI: 10.1525/elementa.379.

Clearstone Engineering and Carleton University, 2018. Update of equipment, component and fugitive emission factors for Alberta Upstream Oil and Gas. 157 pp.

Fox, T.A., Barchyn, T.E., Risk, D., Ravikumar, A.P., Hugenholtz, C.H., 2019a. A review of close-range and screening technologies for mitigating fugitive methane emissions in upstream oil and gas. *Environmental Research Letters* 14, 053002. DOI: 10.1088/1748-9326/ab0cc3.

Fox, T.A., Ravikumar, A.P., Hugenholtz, C.H., Zimmerle, D., Barchyn, T.E., Johnson, M.R., Lyon, D., Taylor, T., 2019b. A methane emissions reduction equivalence framework for alternative leak detection and repair programs. *Elementa: Science of the Anthropocene* 7, 30. DOI: 10.1525/elementa.369.

Ravikumar, A.P., Wang, J., McGuire, M., Bell, C.S., Zimmerle, D., Brandt, A.R., 2018. "Good versus good enough?" Empirical tests of methane leak detection sensitivity of a commercial infrared infrared camera. *Environmental Science and Technology* 52, 2368-2374. DOI: 10.1021/acs.est.7b04945.

Ravikumar, A.P., Sreedhara, S., Wang, J., Englander, J., Roda-Stuart, D., Bell, C., Zimmerle, D., Lyon, D., Mogstad, I., Ratner, B., Brandt, A.R., 2019. Single-blind inter-comparison of methane detection technologies – results from the Stanford/EDF Mobile Monitoring Challenge. *Elementa: Science of the Anthropocene* 7, 37. DOI: 10.1525/elementa.373.

Zavala-Araiza, D., Herndon, S.C., Roscioli, J.R., Yacovitch, T.I., Johnson, M.R., Tyner, D.R., Omara, M., Knighton, B., 2018. Methane emissions from oil and gas production sites in Alberta, Canada. *Elementa: Science of the Anthropocene* 6, 27. DOI: https://doi.org/10.1525/elementa.284.