Detection Limits of Optical Gas Imaging for Natural Gas Leak Detection in Realistic Controlled Conditions

Daniel Zimmerle[1], **Timothy Vaughn**[1], **Clay Bell**[1], **Kristine Bennett**[1], **Parik Deshmukh**[2], **Eben Thoma**[3]

[1]Energy Institute, Colorado State University, Fort Collins, Colorado 80523, United States.

[2]Jacobs Technology, Durham, North Carolina 27712, United States.

[3]EPA Office of Research and Development, Durham North Carolina 27711, United States.

Abstract

Optical gas imaging (OGI) is a commonly utilized leak detection method in the upstream and midstream sectors of the U.S. natural gas industry. This study characterized the detection efficacy of OGI surveyors, using their own cameras and protocols, with controlled releases in an 8-acre outdoor facility that closely resembles upstream natural gas field operations. Professional surveyors from 16 oil and gas companies and 8 regulatory agencies participated, completing 488 tests over a 10 month period. Detection rates were significantly lower than prior studies focused on camera performance. The leak size required to achieve a 90% probability-of-detection in this study is an order-of-magnitude larger than prior studies. Study results indicate that OGI survey experience significantly impacts leak detection rate: Surveyors from operators/contractors who had surveyed more than 551 sites prior to testing detected 1.7 (1.5–1.8) times more leaks than surveyors who had completed fewer surveys. Highly experienced surveyors adjust their survey speed, examine components from multiple viewpoints, and make other adjustments that improve their leak detection rate, indicating that modifications of survey protocols and targeted training could improve leak detection rates overall.

Introduction

Driven by improvements in production methods, the production and use of natural gas in the U.S. has grown rapidly.(1) While natural gas produces lower combustion carbon dioxide emissions than coal, methane, the primary constituent of natural gas, has a higher global warming potential ($\approx$84 times CO_2 over 20 years) leading to an interest in reducing CH_4 emissions,(2) and therefore in natural gas leak detection methods. This study presents recent tests of the detection efficacy of one of the most popular leak detection technologies used in mid- and upstream sectors of the natural gas industry: optical gas imaging (OGI). OGI surveys have also been extensively utilized for leak detection in recent studies of production, (3,4) gathering,(5,6) transmission,(7,8) and distribution systems.(9,10) Therefore, the efficacy of the method is of both policy and research interest.

Emissions of CH_4 and volatile organic compounds (VOCs) from leaking industrial process equipment have traditionally been controlled through leak detection and repair (LDAR) programs using U.S. Environmental Protection Agency (EPA) Method 21 (M21), which

employs a portable hydrocarbon (HC) probe to directly inspect each component.(11) M21 is sensitive and relatively precise, but is resource intensive since the leak surveyor must individually document, physically touch, and thoroughly inspect the entire leak interface of each component. Next generation approaches, such as OGI, operate some distance away from the inspected component (without physical contact) and offer advantages over M21 regarding ease of use and higher spatiotemporal coverage, while still obtaining component-level information. However, OGI, like other next-generation methods, is also affected by a much larger range of factors than M21, such as environmental conditions, operating distance, and interfering sources. OGI is the only EPA-approved alternative to M21 for use in refineries,(12) and is the primary LDAR method for management of VOC leaks (CH_4 is treated as a coemitted gas) in U.S. oil and gas (O&G).(13)

This study focuses on the dominant type of OGI system used in O&G LDAR applications: A video camera using a cooled detector filtered to the 3.2 to 3.4 μm spectra band where normally invisible HC emissions absorb light.(14–16) In practice, both the HC emission plume and the background scene (e.g., equipment, sky, or ground) absorb and emit infrared radiation in the spectral window of the camera. If enough image contrast exists between the HC plume and the background scene, the emission can be detected by the OGI surveyor.(17) While other OGI cameras are available, all surveyors in this study used the FLIR GF320 (GF320) camera (FLIR Inc., North Billerica, MA).

OGI leak detection capability is affected by instrument performance, environmental use conditions, and human factors, with the last being the primary focus of this study. OGI leak detection sensitivity cannot be expressed by a single number, but creates a detection threshold (DT) band, within which there is a variable probability of detection. Instrument detection limit studies in controlled conditions show that OGI is capable of detecting CH_4 leaks in the 1–4 g·hr^{-1} range, with many HCs exhibiting better detection sensitivity than CH_4 due to higher integrated absorption strength in the observed spectral band.(14–21) Prior studies have identified several physical and environmental factors that affect the active DT and can result degradation of detection sensitivity by orders of magnitude. Observation distance, atmospheric dispersion of the emitted plume (e.g., wind speed), and the apparent temperature difference between the emitted gas and background scene (T), which includes both ambient temperature blackbody radiation and direct solar reflection, are significant determining factors in OGI detection performance.(17–23) Whereas instrument and environmental factors affecting OGI detection sensitivities are becoming better understood, human and method protocol effects, such as the difference between proximate high-dwell observations of specific components versus more routine LDAR scanning leaks surveys are just beginning to be studied.(24–27) A limited number of studies have compared OGI, M21-type detections, and emission rate measurements on O&G sites, documenting the sensitivity advantage of hand-held M21 probes in detection of small emissions but also noting the fundamentally different aspects of the approaches and the types of emissions found. Previous field work on O&G sites indicates the ability of OGI to detect small emissions (e.g., <1.5 standard cubic feet per hour (scfh) or ≈30 g·h^{-1}) under favorable conditions, but also to miss larger emissions under unfavorable conditions.(24–27)

In addition to the detection capability of the camera (e.g., noise floor), cameras include software enhancements that assist human operators by increasing the contrast of the emission plume. The GF320, for example, can be operated in a normal (auto adjusting) gray scale imaging mode—where the equipment imaged in the scene is easily recognized—or in a high sensitivity mode (HSM) that uses image differencing to enhance the visibility of small emissions. The HSM is less intuitive but is often preferred by experienced surveyors. Although the HSM can enhance detections, it can also increase the likelihood of false positives, and may have difficulty with scenes that include a large temperature range (Supporting Information (SI) Videos 1 and 5).

The OGI detection threshold (DT) band is also impacted by the surveyor's experience, training, facility conditions, and weather conditions. OGI surveys are conducted in a wide variety of facilities, ranging from well pads to gas processing plants, with component counts ranging across several orders of magnitude. ("Components" includes all constituent parts at a facility, including valves, piping couplers, controllers, actuators, etc.) While protocols vary between companies, a typical protocol requires a surveyor to image each component to look for a visible plume.

Prior studies have focused on camera performance, typically using tripod-mounted cameras and controlled conditions for testing.(16,17,22,23) The authors know of no peer-reviewed studies that looked systematically at DT variations of the combined camera-surveyor system in realistic conditions where leak locations and rates were known and could be tested across a wide range of weather conditions. To fill this gap, this study utilized a single-blind test wherein surveyors, using their own camera and survey protocols, detected leaks on a simulated upstream natural gas facility with a wide selection of controlled emission locations. The study spanned 10 months and included a wide range of weather conditions.

Methods

Test Facility

All testing was performed at theMethane Emissions Technology Evaluation Center (METEC) at Colorado State University (CSU) in Fort Collins, CO. For this study, active METEC equipment was grouped into three test well pads (SI Figure S-1). O&G equipment on the site was assembled from recently decommissioned equipment that was reassembled at METEC and augmented with gas supply lines to simulate leaks. Where possible, gas delivery tubing was hidden from view, or routed in existing small-diameter piping characteristic of the equipment when it was in normal field service–for example, gas delivery lines for pneumatic controllers and actuators. The routing of gas from the supply tanks to emission points was fully hidden on Pad 3 and mostly hidden on Pads 1 and 2.

Gas release rates are set by solenoid valves in series with choked-flow orifices (SI Figure S-2). After site configuration, all emissions were controlled remotely from the onsite control room. Emission rates were measured using Omega FMA-1700 thermal mass flow meters (SI Section S-2.1). Emission rates and enviornmental data were recorded at 1 Hz. Reported emission rates utilized for this study were averaged over the test duration.

The resulting equipment closely resembled functioning field equipment. Emission points were in locations similar to field equipment, and emission rates and patterns were similar to field observations from prior studies.(3,4,28) However, there were several important differences between the configuration of METEC for this study and conditions in an active O&G production basin.

- No equipment was heated by the gas burners attached to the equipment, unlike field conditions, where separators may be heated for process reasons. Heated equipment may provide more background contrast than unheated equipment in some conditions, and leaks on heated equipment may release heated vapors that are more visible in an OGI camera against cold backgrounds.

- In some rare field conditions, leaking gas may be emitted at high pressure and velocity, forming a small jet near the point of the leak. For this study, all gas was emitted at near atmospheric pressure and no jets were formed. While jets are smaller, making them harder to detect, expansion cooling may increase the thermal contrast versus background temperatures.

- To avoid "tipping off" surveyors to the presence of an emission source, industrial-grade, nonodorized, CH_4, was utilized for most tests (SI Section S-3). In some fields, gas contains VOCs which raises plume visibility in OGI cameras. Additionally, in fields with significant oil or condensate production, produced gas has a noticeable odor. Both visibility and increased odor increase the potential for detecting an emitter.

- Human factors differed from field operations. Notably, when surveying at METEC, surveyors tended to be strongly focused on their detection tasks, and typically exhibited a competitive spirit to detect as many leaks as possible, in an environment where they *knew* there would be leaks. In field conditions, surveyors may be less motivated or more distracted, which could lead to different effective performance.

- METEC contains only well pad equipment, while OGI is also utilized on more complex facilities, with more closely packed components, higher noise levels, and more vibration, that make leaks more difficult to detect.

The impact of the difference between METEC and field conditions were impossible to disentangle from the results. It is the judgment of the study team that the factors noted above tend to offset each other—some making detection easier, others making detection harder— and results are strongly indicative of field conditions. However, subsequent users of the data and conclusions should be be aware of the test conditions and potential limitations on applicability.

Every METEC emission point was assigned a unique identifier. SI Table S-4 provides a list of all emission points utilized in the study, including typical viewing direction and background during a leak survey. Accompanying data files include a short description of the leak point, such as the device type and leak location. In general, a single leak point may be utilized in more than one test configuration.

Experimental Design

All surveyors included in this paper were recruited from one of three groups:

- *Compliance* surveyors who work for federal, state, county, or provincial regulators (hereafter, Canadian provincial surveyors are included in the "state" classification). These surveyors screen facilities to ensure compliance with relevant regulations.

- *Company* surveyors who work for O&G operators. All surveyors participating in this study complete periodic leak detection surveys on their facilities; some perform mandatory surveys required for regulatory reporting (e.g., the EPA Greenhouse Gas Reporting Program (GHGRP)).

- *Contractor* surveyors who work for service companies that contract with O&G operators to perform surveys for LDAR or reporting programs.

With the exception of one surveyor, who worked for a contract company that was paid to participate, all surveyors were volunteered by, and funded by, their organizations for the time and travel costs to participate in the study.

In practice, protocols differed more between compliance surveyors and either company or contractor surveyors. Therefore, it is convenient to classify surveyors working for O&G operators, and contractors hired by those operators, as LDAR surveyors—i.e., all surveyors employed by or for O&G companies, as distinct from *compliance* surveyors, who all worked for some public jurisdiction. In total there were 26 LDAR surveyors (23 from O&G operators and 3 from contractors), and 13 compliance surveyors (1 county, 10 state, and 2 federal).

Surveyors were recruited by contacting the organizations where they were employed. CSU directly contacted most LDAR participants, while staff at the EPA contacted compliance surveyors. The *Environmental Partnership*,(29) an industry group, also assisted in recruiting surveyors from their member companies. When any surveyors volunteered to participate, they were scheduled into the available test days that fit their schedules. Between 1 and 4 surveyors were on site at simultaneously, and for approximately one-half of all surveyor-days, the surveyor was accompanied by a second person who helped record notes while the surveyor performed the survey (SI Figure S-3). The presence of an assistant likely reduced the time to complete the survey by off-loading recording work from the surveyor, but is unlikely to have changed the overall detection rate, as the assistant could not see or smell the leak.

Testing was designed to collect leak detection performance data using the same methods surveyors would use in field conditions. Each surveyor brought their own OGI camera and performed the survey using their normal working methodology. The study intentionally allowed the leak detection protocols to vary between surveyors; the intent was to capture, as closely as possible, performance typical of field surveys.

To expand the conditions experienced by the surveyors, testing was conducted between 28 January and 16 November, 2018 (SI Figure S-3). Teams experienced realistic variations in

weather conditions, ranging from cold winter to hot summer days with varying cloud cover, solar insolation, and precipitation conditions. Wind varied between calm and 12 m/s (SI Figure S-4) and the temperature varied between –2 and 32 °C. While surveys were not conducted in unsafe weather, such as thunderstorms or heavy precipitation, surveys *were* conducted in conditions outside the acceptable range for survey protocols to understand the impact on detection rates. Of 162 test configurations completed by surveyors whose protocols had specific numerical wind speed cutoffs, 4 test configurations (2.5%) exceeded the wind cutoff limit.

Teams arrived at METEC with their equipment in the morning and received safety and project methodology briefings. Each surveyor completed a survey form collecting information about their OGI survey experience and training prior to the study, their camera, and their leak detection protocol (SI Section S-2.2).

For the remainder of the day, the surveyors were tasked with finding leaks on the METEC equipment. Testing proceeded in the following sequence: First, a leak pattern was initiated on all three well pads shown in SI Figure S-1, using a preplanned test that set the leak locations, rates, and intermittency. The number of leaks per well pad varied from zero (no leaks) to 7 leaks per pad. Surveyors did not know the location, size, or emission pattern of the leaks. However, for leaks above 3 slpm (≈6 scfh), surveyors could possibly hear the flow of gas through the gas release equipment and/or the leak location, which is similar to field conditions.

When the leak patterns were operating and flow rates calibrated, surveyors were instructed to circulate through the facility and survey using the normal methods. When multiple surveyors were screening simultaneously, they were kept separated from each other (i.e., on separate pads or, in the case of Pad 3, on separate halves of the pad) and instructed to not communicate while working. Surveyors recorded each leak found, including location, viewing distances, and other information (SI Section S-2.3).

When a surveyor finished with one test configuration on all pads, they returned to the control center, finalized their reporting log, and submitted it to METEC staff. Finally, when all surveyors completed leak detection on a test configuration, METEC staff initiated a new test configuration and again dispatched surveyors to detect leaks. Using this method, each surveyor completed 1–4 rounds of the three test pads during each day of testing.

In general, survey protocols varied more between major classifications (i.e., LDAR or compliance) than between surveyors within one classification. For example, LDAR surveyors typically ascend catwalk stairs to inspect tanks and open doors to inspect separator control pneumatics, while many compliance surveyors are instructed to not climb on or open any facility equipment. The protocol for some compliance inspectors was to conduct the survey from outside of spill containment berms–a substantial distance from many components, raising the DT and reducing viewing angle flexibility.

The study included nine test IDs, where each test ID represented one configuration of all three pads at METEC. While the nominal release rate for each leak in each test ID was set a priori, these rates were intentionally varied over the duration of the study to increase the

variability in test conditions (SI Figure S-8). Intermittent releases were included in the tests and the average release rates for these emitters were generally much higher than that of steady-state emitters. Since only a small number of tests included intermittent releases, these were eliminated from most analyses included here; charts note whether intermittent emitters are included in the chart content.

In test ID 1, two release points were intentionally placed close to each other. In postcampaign analysis, however, it proved difficult to separate detections on these components in surveyor detection logs. In addition, since surveyors were not allowed to use soap bubbles to confirm their leaks (to avoid tipping off other surveyors), many surveyors indicated that they would not be able to isolate closely spaced leaks. Therefore, in the analysis, *any* detection proximate to either leak point was included as *one* successful detection on *one* emission point.

Leak locations utilized in this testing were taken from prior field studies, including studies by CSU(28) and other upstream studies.(3) While leak rates are similar to those observed in studies, testing emphasized lower emission rates in an effort to clearly identify emission rates at which detection rates fall off (SI Figure S-9).

Analysis Methods

For this study, we define a *test configuration* as one test ID, on one pad (see SI Figure S-1), surveyed by one operator on 1 day; there were 488 such test configurations in this study, identified by the "ConfigID" field in the data tables. Each is roughly equivalent to a single well-pad survey in field conditions. For each test configuration, each pad was configured in one of two scenarios: (1) a *zero emission* configuration, where no gas was released, and (2) a *nonzero emission* configuration, where gas was released from one or more locations. The 488 total test configurations consisted of 73 zero emission configurations and 415 nonzero emission configurations.

The difference in complexity of the pads is reflected in the time surveyors spent on each. Survey times for Pad 1 (mean: 19 min, range: 3–52 min) and Pad 2 (mean: 18 min, range: 1–89 min) are similar, while surveys on Pad 3 (mean: 39 min, range: 9–108 min) took twice as long on average (SI Figure S-10).

For each pad-level survey, the surveyor could: (a) identify that a well pad had no emissions, or (b) detect apparent leak(s). For the zero-emission tests, the surveyor could correctly identify the pad as nonemitting, i.e., a true negative (TN), or incorrectly identify one or more emissions on the pad, each defined as a false positive (FP). For nonzero emission tests, the surveyor could successfully identify an individual emission point—a true positive (TP), not find an emission point—a false negative (FN), or identify an emission point where none existed—a false positive (FP). Using this common definition, TN is a measure of performance while screening an entire pad, while TP, FN, and FP are measures of performance at an individual leak detection. This study produced 770 TP, 595 FN, 73 TN, 31 FP reports. Reports were coded into a single data table, included in the attached data file (SI Section S-2.4).

To create detection curves by any independent variable, the data were binned by that variable and all test outcomes within each bin are utilized to calculate detection fraction. Confidence intervals (CIs), were calculated using bootstrap methods as discussed in SI Section S-2.5. In keeping with most detection studies, central 90% confidence intervals were used throughout. CIs are expressed throughout as μ (*l* to *u*), where μ is the mean value, and *l* and *u* are the lower and upper confidence bounds, respectively. To distinguish "changes between two conditions" from confidence intervals, the word "from" is added: " (from *s* to *e*)", where is the net movement, *s* is the starting value, and *e* is the ending value (e.g., −10% (from 22% to 12%)). Emission rates are expressed two ways: in slpm, with standard conditions of 20 °C and 1 atm, and in scfh, with the standard conditions at 70 °F, 14.7 psia, leading to a conversion factor of 0.47 slpm/scfh. For emissions of industrial CH_4, 1 scfh ≈18.8 g/h.

The influence of independent variables, such as emission rate, wind speed, ambient temperature, surveyor training, and experience, etc., on the detection rate was the primary interest of this study. However, since independent variables may be correlated with each other, attributing results to specific inputs must be done with caution. For example, wind speed is known to impact detectability of emissions, and a reasonable hypothesis is that surveyors may tire late in the day and become less effective. However, since winds tend to be higher in mid-to-late afternoon (SI Figure S-5), the impact of these two variables cannot be completely separated.

Results and Discussion

Demographic Data

A total of 39 surveyors (26 LDAR and 13 compliance) from 24 organizations participated in the study over 26 days of testing. Individual organizations contributed 1–3 surveyors. Surveyors participated in an average of 4.1 test configurations (range of 1–8). In general, surveyors completed each pad-level test once. However, 18 test configurations were repeated twice, involving five surveyors (BB1, UD1, UD2, ZG1, and ZG2) and three test IDs (T2, T5, and T6).

Surveyors self-reported metadata about their training, cameras, and prior experience (SI Section S-2.2). Most protocols required surveyors to be trained on OGI or thermography; 90% reported having formal training in OGI, and of those with training, 8.6% had taken the more intensive "Level 2" training. Prior to the study, surveyors reported having screened between 1 and ≈4000 facilities; 7 had surveyed fewer than 100 facilities, while 11 reported surveying at least 1000 facilities.

Surveyors followed formal and informal protocols while surveying. Some surveyors provided copies of their organization's protocol documents, while others did not provide a written protocol or indicated that no such document existed. One key guideline in most protocols was the maximum wind speed under which a survey can be conducted, leaks disperse more in higher winds and are harder to detect. Of 39 surveyors, 17 reported a specific wind speed cutoff, ranging from 4.5 to 16 m/s. Using the average of these cutoffs as a guide (9.3 m/s, or 21 miles/hour), we limit most analyses to tests performed at wind speeds

below 9 m/s, which includes 89% of all tests. For the 47 tests performed with wind speeds above the cutoff, leaks averaged 7.5 (0.06 to 30) scfh and 51% were detected.

Detection Rate for Nonzero Emissions

The detection rate of nonzero emissions represents the primary measure of efficacy; i.e., what fraction of existing emissions are detected by surveyors? Of all independent variables captured, surveyor experience and affiliation had the largest impact on detection rate. (Other key variables are analyzed in SI Section S-7). To analyze experience, surveyors were divided into groups based upon the self-reported number of facilities they had screened prior to the study. For most analyses, two compliance surveyors (WV1 and QO1), who had surveyed <10 sites prior to the study, were not included.

The cutoff between low- and high-experience groups was selected to qualitatively maximize the contrast between experience categories. Data indicated that experience not only has a dramatic impact on detection rate, but the difference above/below 551 surveys is statistically significant and robust across a wide range of possible cutoffs (SI Section S-5). In addition, no compliance surveyor had surveyed more than 550 sites prior to the study, while multiple LDAR surveyors reported more than 700 sites (SI Figure S-11). The results of a sensitivity analysis also indicated similar detection rates across all compliance surveyors, but substantially different detection rates for low- and high-experienced LDAR surveyors (SI Section S-5). Therefore, surveyors were divided into three groups, low (25–200 sites) and high experience (700–4000 sites) LDAR surveyors, and all compliance surveyors with some experience (12–550 sites).

Figure 1 illustrates the significant increase in detection rate for surveyors who had completed more than 551 surveys prior to the study. Considering both LDAR and compliance surveyors, surveyors in the high experience group find 1.6 (1.5–1.7) as many leaks as those with less experience; this increases to 1.7 (1.5–1.8) when considering only LDAR surveyors. Further, the binned data in both groups show strong self-similarity, with a marked difference between the groups. For example, 11 of 12 (92%) high-experience LDAR surveyors had individual detection rates 60%, while 3 of 10 (30%) low-experience LDAR, and 3 of 13 (23%) compliance, surveyors exceeded 60% detection rate. Additional analysis indicates that moving the cutoff over a 2:1 range changes detection rate by 5.8 percentage points in any experience category, in part because there is a significant performance gap between low- and high-experience LDAR surveyors (SI Section S-5).

Emission rate also has a primary impact on detection rate–higher emission rates create higher gas concentrations, which are generally easier to detect. Figure 2 illustrates detection curves by emission rate, segmented by experience and affiliation. The detection rates seen in this study are substantially lower than prior camera-focused studies, which did not include surveyor performance in testing.(16–18,23) Considering the group with the highest detection rates, high-experience LDAR surveyors, the estimated 90% probability-of-detection is 3.29 (2.6–7.7) slpm (7 (5.52–16.4) scfh). In contrast, at this study's mean detection distance of 2.7 m (8.9 ft), Ravikumar et al.(23) estimated a 90% probability-of-detection at 13 g/h (0.7 scfh), an order of magnitude lower emission rate. In addition, low-experience and compliance surveyors have substantially lower detection rates (SI Section S-4).

While direct comparison between the studies is difficult due to the different foci and methods, this pronounced difference indicates that the *camera-surveyor system*, working across a wide, but realistic, range of environmental conditions, attempting to find unknown leak locations throughout the facility, had a substantially lower DT band than the camera itself, viewing a known leak location from a tripod.

Using the above experience bins, the impact of experience on detection rate was significantly different between high-experience LDAR and either low-experience LDAR or compliance surveyors. High-experience LDAR surveyors (700–4000 sites) had a detection rate that was 1.7 (1.5–1.8) times that of low experience LDAR surveyors (75% (72–78%) versus 45% (41–49%)). Additionally, the detection rate for compliance surveyors is statistically similar to that of low-experience LDAR surveyors (49% (44–54%)) (SI Figure S-13). The similarity between these two groups is likely due to protocol differences between compliance and LDAR surveys (SI Section S-6), and the substantial overlap in experience between compliance and low-experience LDAR surveyors (SI Figure S-11).

Additional data collected during the campaign provided some indications of *why* more experienced surveyors have higher detection rates. Two variables, inspection times and viewing background, are of particular interest; others are given in SI Section S-5.

Inspection Times

To analyze inspection times, time spent by surveyors while performing inspections was normalized to the number of "major equipment units" (a well head, a tank or a separator) on each well pad for each test (Figure 3a). Detection rates generally fell as inspections speeds increased. Survey times of <2 min/unit produced statistically lower detection rates than surveys of 3 min or more for both low-experience LDAR surveyors and compliance surveyors. These differences were substantial: Increasing survey times from 2 min to >3 min increased detection rates by 1.3 times for high-experience LDAR surveyors, by 1.8 for compliance surveyors and by 2.4 times for low-experience LDAR surveyors.

Additionally, experienced LDAR surveyors spent more time inspecting than the other surveyor groups (Figure 3b). Highly experienced LDAR surveyors screened 18% of major equipment units in <2 min, while spending at least 3 min on 39% of units. Less experienced LDAR surveyors and compliance surveyors effectively inverted these speeds, screening 46% of units in <2 min, and spending more than 3 min on 24% of units.

While it is not clear if these speed differences would be present in field surveys as well, these data indicate that minimum inspection times should be considered for survey protocols, and recommended speeds should potentially be conditioned on surveyor's experience. METEC has only a small range of equipment complexity—well heads, separators, and tanks—but for this range of complexity, inspection times of >3 min/unit are recommended for most units, as practiced by highly experienced LDAR surveyors.

Background View

The background behind the leak location also has an impact on detecting a leak. Two factors are in play. First, the background may have a similar temperature as released gas, reducing

contrast between the plume and background in the video image. Second, the background may have moving or flickering elements, such as plants moving in a breeze or heat shimmering from a road surface. This visual noise imparts motion to the scene that may make the plume harder to detect.

To understand the impact of background, the scene behind each emission point was identified after the campaign: Each leak location was viewed as it would commonly be viewed during inspections, and the background and angle of view was recorded. If a leak location could be naturally inspected from more than one direction, and if those directions had different backgrounds, a secondary background was also recorded (SI data file D2 *LeakPointData*, SI Table S-2). For this analysis we use only the primary background classification.

Three background types were identified: (a) "ground" indicates that the leak was viewed against the ground, either near the leak location or at a distance. Since many leaks are below eye level for most surveyors, 48% of leaks fall into this category; (b) "equipment" indicates the background was other equipment, such as the surface of a separator or tank, that was <1 m away (30% of tests); and (c) "sky" indicates that the leak was likely viewed against a sky background (21% of tests).

Emissions simulated in the study utilized similar rates as those found in field studies. As a result, emissions from tanks, which are more likely to viewed against the sky, tended to be larger than other leak locations, a mean of 9.6 scfh for sky backgrounds versus 3.1 for ground backgrounds and 3.3 scfh for equipment backgrounds. To avoid bias, this analysis considers only emissions 8 scfh, which reduces the mean emission rate for sky backgrounds to 3.4 scfh.

The scene behind components being inspected had a substantial impact on detection rates. When leaks were viewed against the ground, there was a significant decrease in detection rates relative to sky or equipment backgrounds, regardless of experience. For the high-experience surveyors, the change in detection rate was −10% (from 75% to 65%) when viewing potential leaks against the ground rather than the sky. The difference was larger for surveyors with less experience: −17% (from 46% to 29%) sky-to-ground and −32% (from 61% to 29%) equipment-to-ground. SI Video 2 provides an example of a plume that is/is not visible against the ground, while remaining visible against nearby equipment (SI Section S-8.2), SI Video 3 illustrates the difference in visibility for sky and equipment backgrounds (SI Section S-8.3), and SI Video 4 illustrates the difference in visibility inside a separator cabinet in three viewing configurations (SI Section S-8.4).

These data suggest that survey protocols should advise additional care when viewing a component against a ground background. This would likely impact survey time or procedures for (a) components that are substantially below eye level, such as most well head components or drain equipment on separators, and (b) components higher off the ground where there is no nearby equipment to provide a background for inspection. The data and videos also indicate that components should be viewed against multiple backgrounds if at all possible to improve detection rates. See Figure 4.

False Positives

False positives occur in two contexts–zero emission pad-level tests and as additional mistaken detections on pads with one or more emitters.

For the 73 "zero emission" pad-level tests, *any* false positives on the pad are counted as a single false positive, i.e., a measure of how often surveyors indicate that a test pad with no emitters has emissions. Overall, the false positive rate was 4.1% (0–7.9%). Excluding surveyors who had surveyed <10 sites prior to the study, high experience surveyors reported false positives in 1 of 30 tests (3.3% (0–9.7%)), whereas those with less experience reported false positives in 1 of 23 tests (4.2% (0–13%)); compliance surveyors had no false positives on this type of test (0 of 15 tests). These data indicate that the performance differences by experience do not extend to false positive rate; all groups are similar.

The second type of false positive is a mistaken report of emissions on a pad with other emission points. In this case, false positives provide an indication of the number of false detects as a fraction of the total number of leak reports. For this type of false positive, more experienced surveyors had a higher false positive rate (9 of 490 tests (1.8% (0.8–2.8%))) than less experienced surveyors (1 of 453 tests (0.22% (0–0.66%))). These data indicate that the higher detection rate of high-experience surveyors came at the price of slightly higher false positive rate.

For inexperienced surveyors, (<10 sites screened prior to the study), false positive rates increase dramatically. For zero emission testing surveyors had a false positive rate of 1 of 5 tests (17% (0–50%)); for tests with emissions, they had a false positive rate of 9 of 89 tests (9.4% (5.1–14%)).

Guidance

The difference between this study and prior studies is that prior studies (with the exception of Pacsi et al.(27)) focused on the sensitivity of the camera, while this study (a) tested the camera-surveyor system, (b) using the surveyor's protocol, (c) while detecting common leak sizes in real time on realistic equipment in a wide range of weather conditions, (d) in a controlled facility with knowledge of every emission. Results indicate overall detection rates from the camera-surveyor system are substantially lower than the detection rates estimated from studies focused solely on the camera performance(16–18,22,23) or field studies where ground truth emissions were only partially known.(27) Given the complexity of the survey task, this is unsurprising. Our data also indicate that all surveyors have a high probability of detecting large leaks, those above 10 slpm (21 scfh) under favorable meteorological conditions (SI Section S-4). These results indicate that emitter counts from prior studies (including those of the authors(5,6,8,28)) are likely underestimated, while it is unclear if total emissions—which are generally dominated by large emitters—would be highly impacted. For example, although tests performed here were intentionally biased toward low emission rates, surveyors detected 69% of total emissions (SI Figure S-22).

Additionally, these data provide a more robust baseline to compare OGI to next generation leak detection solutions—which are typically automated to a greater degree than OGI

surveys—and data to estimate the overall efficacy of LDAR programs that use OGI as a principal means of leak detection.

While surveyor experience was the strongest predictor of detection rates, differences in protocols between LDAR surveyors (who generally move close to equipment) and compliance surveyors (who generally survey from farther away) also contribute to detection differences (SI Section S-6), which may have made compliance surveyor detection rates substantially similar to low-experience LDAR surveyors. These data indicate that researchers should be conscious of protocol differences when comparing results between studies. For example, compliance surveyors screening well pads with protocols tuned to find open tank hatches, stuck dump valves, or unlit flares, are unlikely to acquire complete counts of all emitters on screened facilities.

The impact of surveyor experience on detection rate indicates a need for additional, or different, training programs for surveyors. Estimating that a surveyor can survey 4–5 well pads per day in a dense field, accumulating 300–400 facilities of experience requires 12–20 weeks of practice with active mentoring. The difficulty with field experience is that the count and location of leaks are unknown, making closed-loop training feedback difficult. Further observations, detailed above, also indicate that behaviors of experienced surveyors could potentially be taught or encoded in protocols, appropriate survey speeds, adjusting location to view components against multiple backgrounds, limiting wind speeds (SI Figure S-20), and similar precautions. Potentially this learning curve could be shortened by training at a facility with known, small, leaks, where directed instruction and immediate feedback could help surveyors gain skills faster.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. EIA. U.S. Natural Gas Marketed Production. https://www.eia.gov/dnav/ng/hist/n9050us2a.htm.

2. Pachauri R; Meyer L; Meyer L IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I II and III to the Fifth Assessment Report of the intergovernmental panel on Climate Change, 2014.

3. Allen DT; Torres VM; Thomas J; Sullivan DW; Harrison M; Hendler A; Herndon SC; Kolb CE; Fraser MP; Hill AD; Lamb BK; Miskimins J; Sawyer RF; Seinfeld JH Measurements of methane emissions at natural gas production sites in the United States. Proc. Natl. Acad. Sci. U. S. A. 2013, 110, 17768–17773, DOI: 10.1073/pnas.1304880110 [PubMed: 24043804]

4. Allen DT; Pacsi AP; Sullivan DW; Zavala-Araiza D; Harrison M; Keen K; Fraser MP; Daniel Hill A; Sawyer RF; Seinfeld JH Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Pneumatic Controllers. Environ. Sci. Technol. 2015, 49, 633–640, DOI: 10.1021/es5040156 [PubMed: 25488196]

5. Vaughn TL; Bell CS; Yacovitch TI; Roscioli JR; Herndon SC; Conley S; Schwietzke S; Heath GA; Ptron G; Zimmerle D Comparing facility-level methane emission rate estimates at natural gas gathering and boosting stations. Elem Sci. Anthropocene 2017, 5. 71 DOI: 10.1525/elementa.257

6. Zimmerle D; Vaughn T; Luck B; Lauderdale T; Keen K; Harrison M; Marchese AJ; Williams LL; Allen DT Methane Emissions from Gathering Compressor Stations in the U.S. Environ. Sci. Technol 2020 54 7552 DOI: 10.1021/acs.est.0c00516 [PubMed: 32407076]

7. Zimmerle DJ; Williams LL; Vaughn TL; Quinn C; Subramanian R; Duggan GP; Willson B; Opsomer JD; Marchese AJ; Martinez DM; Robinson AL Methane Emissions from the Natural Gas Transmission and Storage System in the United States. Environ. Sci. Technol 2015, 49, 9374–9383, DOI: 10.1021/acs.est.5b01669 [PubMed: 26195284]

8. Subramanian R; Williams LL; Vaughn TL; Zimmerle D; Roscioli JR; Herndon SC; Yacovitch TI; Floerchinger C; Tkacik DS; Mitchell AL; Sullivan MR; Dallmann TR; Robinson AL Methane Emissions from Natural Gas Compressor Stations in the Transmission and Storage Sector: Measurements and Comparisons with the EPA Greenhouse Gas Reporting Program Protocol. Environ. Sci. Technol 2015, 49, 3252–3261, DOI: 10.1021/es5060258 [PubMed: 25668051]

9. Lamb BK; Edburg SL; Ferrara TW; Howard T; Harrison MR; Kolb CE; Townsend-Small A; Dyck W; Possolo A; Whetstone JR Direct Measurements Show Decreasing Methane Emissions from Natural Gas Local Distribution Systems in the United States. Environ. Sci. Technol. 2015, 49, 5161–5169, DOI: 10.1021/es505116p [PubMed: 25826444]

10. Zimmerle DJ; Pickering CK; Bell CS; Heath GA; Nummedal D; Pétron G; Vaughn TL Gathering pipeline methane emissions in Fayetteville shale pipelines and scoping guidelines for future pipeline measurement campaigns. Elem. Sci. Anthropocene 2017, 5. 70 DOI: 10.1525/elementa.258

11. U.S. EPA. Determination of Volatile Organic Compound Leaks. 1990

12. U.S. EPA. Alternative work Practice to Detect Leaks from Equipment. 2006, 40 CFR Part 60, [EPA-HQ-OAR-2003–0199; FRL-8055–2], RIN 2060A-L98. 2006

13. U.S. EPA. Oil and Natural Gas Sector: Emission Standards for New, Reconstructed, and Modified Sources Reconsideration. 2018, 40 CFR Part 60, [EPA-HQ-OAR-2017–0483; FRL-9984–43-OAR],RIN 2060-AT54. 2018

14. Footer TU Technical Support Document: Optical Gas Imaging Protocol, 40 CFR Part 60, Appendix K, Revision 5; Technical Support Document EPA-HQ-OAR-2010–0505-4949, 2015; p 255

15. Footer T; DeWees J; Thoma E; Squier B; Eisele A; Secrest C Performance Evaluation and Quality Validation System for Optical Gas Imaging Cameras that Visualize Fugitive Hydrocarbon Gas Emissions. In AWMA 2015 Annual Conference, Raleigh, North Carolina, 2015

16. Zeng Y; Morris J; Sanders A; Mutyala S; Zeng C Methods to determine response factors for infrared gas imagers used as quantitative measurement devices. J. Air Waste Manage. Assoc. 2017, 67, 1180–1191, DOI: 10.1080/10962247.2016.1244130

17. Zeng Y; Morris J Detection limits of optical gas imagers as a function of temperature differential and distance. J. Air Waste Manage. Assoc. 2019, 0, 1–11, DOI: 10.1080/10962247.2018.1540366

18. Benson R; Madding R; Lucier R; Lyons J; Czerepuszko P Standoff Passive Optical Leak Detection of Volatile Organic Compounds using a Cooled InSb Based Infrared Imager. New Orleans, LA, 2006

19. Panek J; Drayton P; Fashimpaur D Controlled Laboratory Sensitivity and Performance Evaluation of Optical Leak Imaging Infrared Cameras for Identifying Alkane, Alkene, and Aromatic Compounds. In Proceedings of the 99th Annual Conference of the Air and Waste Management Association, New Orleans, LA, 20–23 June 2006

20. Sauger E; Fily S; Lejeune H; Thomas B Investigating the use of infrared cameras to detect VOCs. Sealing Technology 2013, 2013, 8–12, tex.ids: saugerInvestigatingUseInfrared2013a DOI: 10.1016/S1350-4789(13)70142-9

21. Hashmonay R; Yanai O; Wahnon A Optical Gas Imaging Sensitivity Study: Envelope of Sensitivity Performance for a Cooled Detector Camera; Opgal: Durham, NC, 2019.

22. Ravikumar AP; Wang J; Brandt AR Are Optical Gas Imaging Technologies Effective For Methane Leak Detection?. Environ. Sci. Technol 2017, 51, 718–724, DOI: 10.1021/acs.est.6b03906 [PubMed: 27936621]

23. Ravikumar AP; Wang J; McGuire M; Bell CS; Zimmerle D; Brandt AR Good versus Good Enough?" Empirical Tests of Methane Leak Detection Sensitivity of a Commercial Infrared Camera. Environ. Sci. Technol 2018, 52, 2368–2374, DOI: 10.1021/acs.est.7b04945 [PubMed: 29351718]

24. Thoma ED; Deshmukh P; Logan R; Stovern M; Dresser C; Brantley HL Assessment of Uinta Basin Oil and Natural Gas Well Pad Pneumatic Controller Emissions. J. Environ. Prot 2017, 8, 394–415, tex.ids: thomaAssessmentUintaBasin2017a DOI: 10.4236/jep.2017.84029

25. Thoma ED Erratum to "Assessment of Uinta Basin Oil and Natural Gas Well Pad Pneumatic Controller Emissions" [Journal of Environmental Protection, 2017, 8, 394–415]. J. Environ. Prot 2017, 08, 1445–1445, tex.ids: thomaErratumAssessmentUinta2017a DOI: 10.4236/jep.2017.812088

26. Stovern M; Murray J; Schwartz C; Beeler C; Thoma ED Understanding oil and gas pneumatic controllers in the DenverJulesburg basin using optical gas imaging. J. Air Waste Manage. Assoc 2020, 70, 468–480, Taylor & Francis _eprint: DOI: 10.1080/10962247.2020.1735576

27. Pacsi AP; Ferrara T; Schwan K; Tupper P; Lev-On M; Smith R; Ritter K Equipment leak detection and quantification at 67 oil and gas sites in the Western United States. Elem Sci. Anth 2019, 7, 29, tex.ids: pacsiEquipmentLeakDetection2019a DOI: 10.1525/elementa.368

28. Bell C; Vaughn T; Zimmerle D; Herndon S; Yacovitch T; Heath G; Ptron G; Edie R; Field R; Murphy S; Robertson A; Soltis J Comparison of methane emission estimates from multiple measurement techniques at natural gas production pads. Elem. Sci. Anthropecene 2017, 5. 79 DOI: 10.1525/elementa.266

29. The Environmental Partnership 2019 Annual Report Summary 2019. 2019.
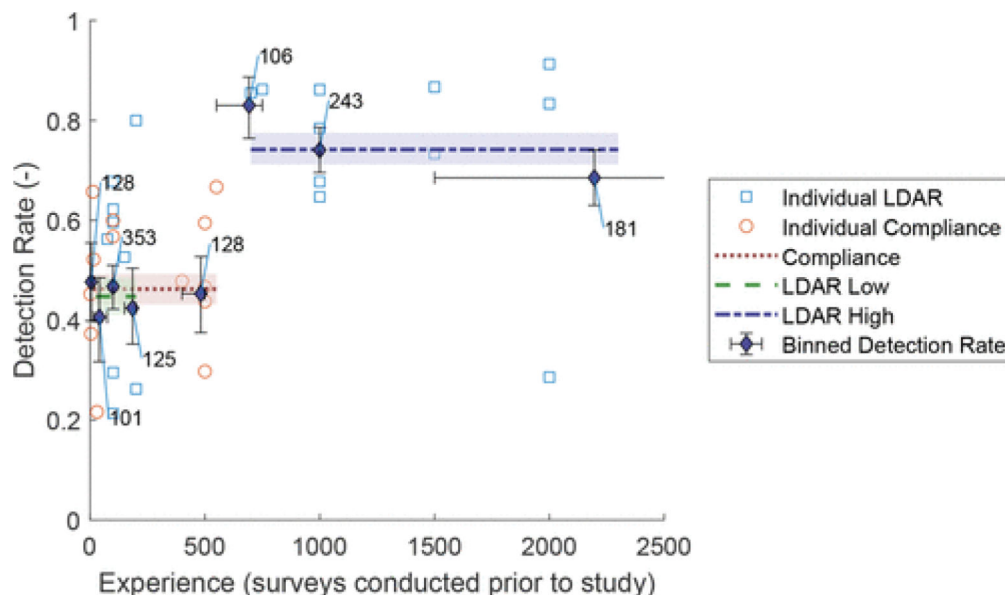
Figure 1.
Impact of experience on detection rate. Filled points indicate detection rate for test data binned by experience, vertical error bars represent bootstrap 90% confidence interval on detection rate and horizontal bars that indicate the range of experience in each bin. Data were binned to ensure 100 tests in each bin, test counts are indicated by numbers. Dotted horizontal lines summarize the mean detection rate each surveyor group used in the analysis; shaded areas are 90% bootstrap confidence intervals. All compliance surveyors had performed less than 551 surveys prior to the study. Unfilled points indicate the detection rate of individual surveyors who had completed more than 15 nonzero emission tests. Range of horizontal axis is reduced from 4000 to 2500 for clarity, clipping off one individual point. Also see SI Table S-7 and SI Figures S-12 and S-13.
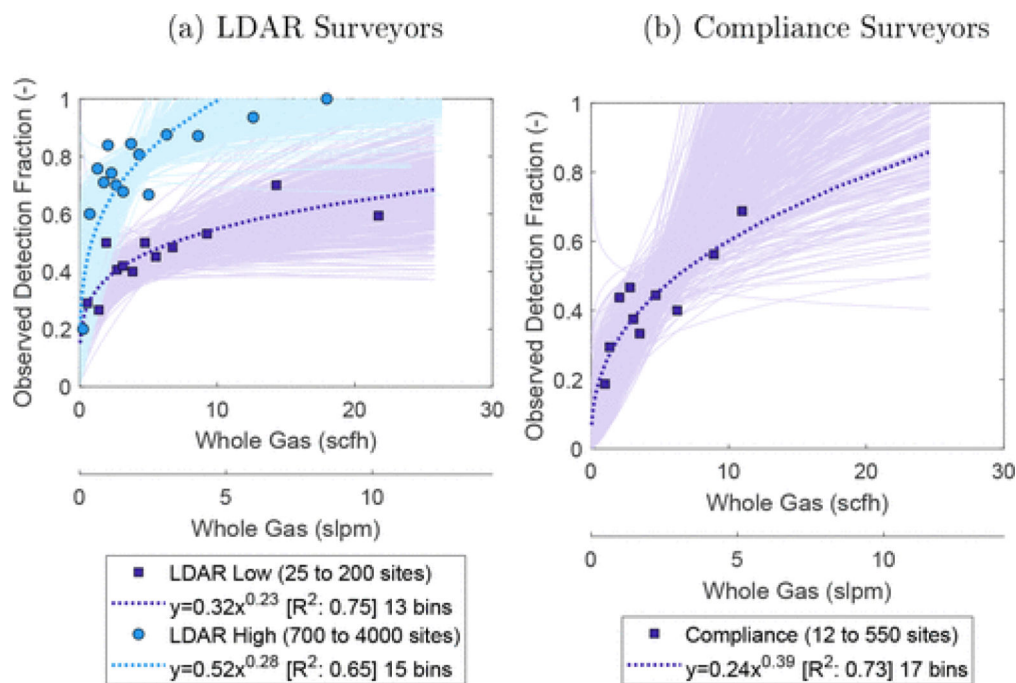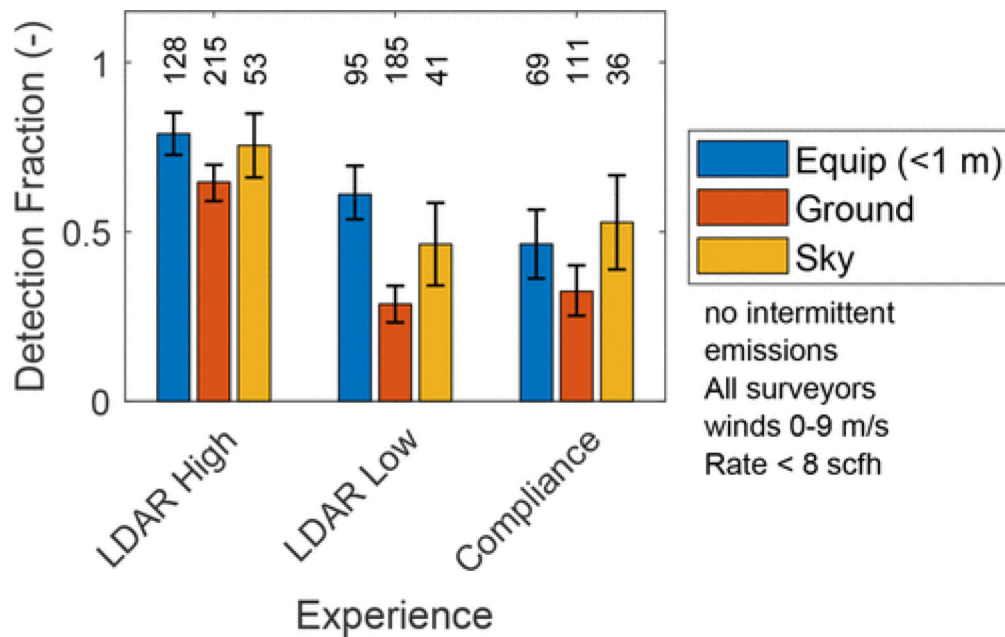
Figure 2.
Overall detection rate for tests with nonzero emissions with outliers of wind speed and temperature excluded. Tests are divided into three groups to illustrate the impact of experience: LDAR surveyors who had surveyed 700 or more sites prior to the study (panel (a), blue), LDAR surveyors who had surveyed 25–200 sites (panel (a), purple), and all compliance surveyors (panel (b)). No LDAR surveyors had prior experience between 200 and 700 sites. Shadow lines are 500 bootstrapped curve fits of the input data. Bin size was selected to produce a similar number of bins for both affiliation groups (15 for compliance, 30 for LDAR). The experience range of compliance surveyors was compressed relative to LDAR surveyors, compliance survey protocols utilize different inspection distances (see text), and compliance surveyors tested over a narrower range of wind conditions (SI Figure S-14). Plots are restricted to winds 0–9 m/s, and ambient temperature 0–35 C, and exclude tests without emissions and all intermittent emitters.
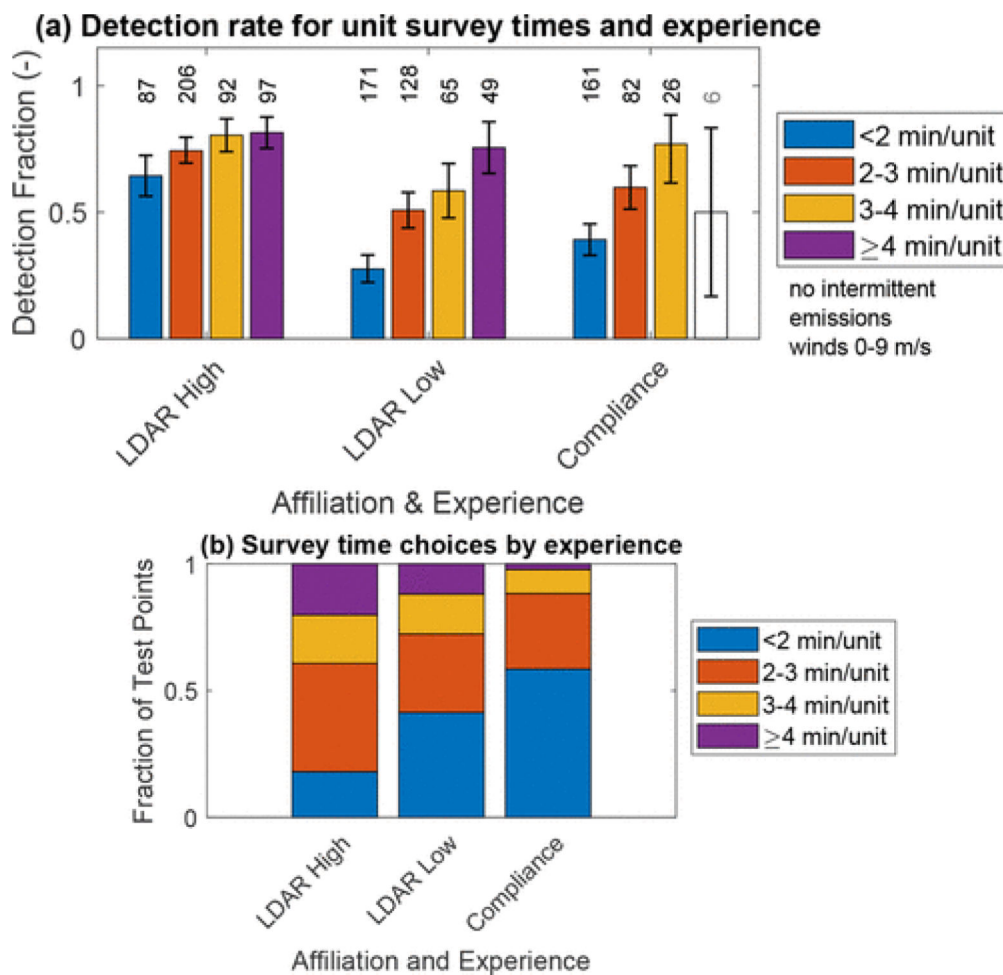
Figure 3.
Time spent inspecting by experience level. LDAR surveyors are split into two experience levels: low (25–200 sites) and high (700 sites), while compliance surveyors are taken as a single group. Time data is normalized to major equipment units, i.e., a well head, tank, or separator. Unshaded bar(s) indicate categories with fewer than 10 data points and insufficient data to draw conclusions. Panel (a): Detection rates generally increase with additional screening time, and for low-experience LDAR surveyors, screening units for <2 min significantly reduces detection rates. Panel (b): High-experience LDAR surveyors screen equipment for longer periods than do either low-experience LDAR or compliance surveyors. The combination of higher screening times and higher detection rates at *any* screening time, raises the detection rates of highly experienced LDAR surveyors.

Figure 4.
Detection rate by background behind leak location. See text for definition of background types. Since leaks with sky background tended to be larger those with other backgrounds, data for this analysis limits includes only test configurations with an emission rate 8 scfh (3.7 slpm). Ground backgrounds have lower detection rates than sky or equipment backgrounds, likely due to the gas and ground having similar temperatures, and the increase visual noise from moving vegetation.