Validation of Remote Sensing and Leak Detection Technologies Under Realistic and Differing Conditions

Contact:

Michael Adamo, PE

Vice President of Operations, OTD

Michael.adamo@otd-co.org

847-544-3428

Prepared by:

Gas Technology Institute

Des Plaines, Illinois

FINAL REPORT

GTI PROJECT NUMBERS 22660 & 22690

Validation of Remote Sensing and Leak Detection Technologies Under Realistic and Differing Conditions

Agreement Number:
693JK31910006

Report Finalized:
March 21, 2022

Prepared For:
U.S. Department of Transportation
Pipeline and Hazardous Materials Safety Administration
Office of Pipeline Safety

Operations Technology Development
Michael Adamo, P.E.
Vice President of Operations

GTI Technical Contact:
Dr. Christopher Moore, Ph.D
Program Manager
847-768-0688
cmoore@gti.energy

1700 S. Mount Prospect Rd.
Des Plaines, Illinois 60018
www.gti.energy

Signature Page

Print or typed First M. Last	*Signature*	*Date*
AUTHOR: **James Taff-Clay**	James Taff-Clay Digitally signed by James Taff-Clay DN: C=US, O=Gas Technology Institute, CN=James Taff-Clay, E=jtaffclay@gti.energy Reason: I am the author of this document Location: Des Plaines,IL Date: 2022.03.23 12:13:24-05'00' Foxit PDF Reader Version: 11.2.1	March 23, 2022
	Title: Associate Engineer	
REVIEWED BY: Chris Moore	Christopher W. Moore Digitally signed by Christopher W. Moore DN: C=US, O=Gas Technology Institute, CN=Christopher W. Moore, E=cmoore@gti.energy Reason: I have reviewed this document Location: Cary, IL Date: 2022.03.22 08:27:11-05'00' Foxit PDF Reader Version: 11.2.1	March 22, 2022
	Title: Program Manager	
APPROVED BY: Amanda Harmon	Amanda Harmon Digitally signed by Amanda Harmon DN: cn=Amanda Harmon, o=GTI, ou=GTI Energy Delivery and Utilization, email=aharmon@gti.energy, c=US Date: 2022.03.23 11:09:10 -05'00'	March 22, 2022
	Title: Manager, Projects	

Legal Notice

Table of Contents

Table of Figures

Page

List of Tables

Executive Summary

The recent and rapid development of new remote sensing-based methane leak detection technologies has prompted a need to properly evaluate and validate those technologies. The new technologies are full systems advanced sensors that contain not only highly sensitive methane analyzers but also a suite of other instruments and algorithms required to remotely locate leaks. Traditional evaluations/validations tended to focus on the methane sensor portion of system instead of the full system. In this project, we developed and tested a method for evaluation of full drone-based methane detection remote sensing systems, with a focus on natural gas transmission pipelines in hard to access areas.

The evaluation focused on a single drone-based system offered by SeekOps, Inc (Austin, TX; SeekOps). A single system was chosen to be able to develop the drone-based system and test that system in a single project. This provided the opportunity to develop an in-depth evaluation method of a single system that could then be used to perform more cursory evaluations of other technologies in the future.

The selected evaluation method involved receiver operating characteristic (ROC) curves and area under the curve (AUC) statistics. These methods quantify tradeoffs between true positives (actual detections) and false positives (detections of leaks that do not exist) and present them as a single number that can be compared across test scenarios and instruments.

Field tests were performed at three locations 1) the Colorado State University Methane Emissions Technology Evaluation Center (METEC), 2) a utility owned natural gas transmission right of way (ROW), and 3) a privately owned ranch near production sites. The SeekOps drone technology showed promise for detecting belowground natural gas transmission pipeline leaks across the range of leak rates studied (0 – 45 scfh) at the three sites. According to our testing, we learned that for the drone to be most effective (i.e., have the highest AUC scores across test scenarios) it should 1) make multiple passes across the area being surveyed and 2) fly in a box pattern around the around being surveyed. When doing these two things the drone system could achieve in some cases AUC scores near 1, indicating a perfect tradeoff between true positives and false positives.

The evaluation testing did identify interesting issues that must be considered when performing evaluation testing and using remote sensing techniques to locate leaks. First, the presence of fog or an atmospheric capping inversion can cause local atmospheric methane concentrations to increase to levels that may make detection impossible at times when in areas of heavy oil and gas activity. Second, rigid survey procedures can limit the ability of the drone system to properly locate leaks, however, structured surveys allow for the proper coverage of all areas to be surveyed. In the real world, to take advantage of the ability of the drone to cover large areas there should be a combination of first flying a predefined path to ensure proper coverage, then targeted investigations with the drone in areas where there were detections. This further highlights the need for a rotary drone for the “hard to access” scenarios studied in this project.

In addition to the methane sensing SeekOps system, a software to identify integrity threats in videos developed by the University of Dayton (UD) was also preliminarily evaluated. The UD integrity threat monitoring software showed promise for being able to locate potential integrity threats among a crowded landscape at METEC. This indicates that given the time to allow the system to train on various images of potential integrity threats (e.g., heavy equipment, field

trucks, people) along natural gas transmission pipeline ROWs, the system could progress into something quite useful.

One area that was not able to be fully examined in this evaluation study was the ability of the drone system to detect leaks that occur on piping submerged in water, including whether the presence of water can impact the way the leak appears at the surface of the water or whether other factors may impact the ability of the drone to detect a submerged leak. Future work should focus on a full evaluation of various leak sizes at different depths below water to explore any impacts that submerged infrastructure may have on the ability of the drone to detect leaks. Also, further development and testing of the drone mounted integrity threat monitoring system is needed.

Introduction

Ensuring the continued safe operation of the 2.2+ million miles of natural gas distribution and 300,000+ miles of transmission pipelines in the U.S. is facilitated by deploying new and innovative technologies. Fast, accurate identification of leaks and monitoring for integrity threats, such as unauthorized 3rd party excavation or encroachment is crucial for reducing overall environmental impacts created by methane releases and reduce risks associated with undetected natural gas leaks. The importance of monitoring the natural gas pipeline system has driven recent technology development specifically using remote sensing instruments on drones.

The advancements in leak detection technologies related to oil and gas started to be described in reports like the Interstate Technology Research Council 2018 overview of existing and emerging leak detection technologies which includes guidance on evaluating leak detection instrument performance[1]. Further, remote sensing technologies mounted to drone platforms were advanced through early programs such as PHMSA[2], DOE ARPA-E MONITOR[3], and the Stanford Natural Gas Initiative/Environmental Defense Fund[4] Mobile Methane Challenge.

Several recent publications have reported on the effectiveness of some of the earlier development activities focused on drone-mounted systems [5-8]. In particular, Barchyn et al 2019[5] tested a fixed-wing drone mounted leak detection system and found that at the higher altitudes used by a fixed-wing drone (40 – 50m), there was a lack of correlation between detection capability and leak size (or rate) for controlled releases. This indicated that drones that fly at lower altitudes may be more effective at locating leaks. The capability for a low flying drone-based leak detection system was further demonstrated by Ravikumar et al. 2019[7]. They showed that a 1 – 3 m altitude from an aboveground, controlled emission, the drone-based leak detection system from SeekOps, Inc. (Austin, TX), was capable of consistently identifying leaks smaller than 1 standard cubic feet per hour (scfh) with near 100% detection probability. The leak detection capability for drone-mounted systems is quite important since drones can be deployed to survey hard to access areas, with low impact to the environment or disruption to the public. There is also a unique co-benefit to leak sensing from drones – the ability to also monitor for ROW integrity threats.

The results from previous studies are promising but were obtained at large field laboratory facilities with leaks released aboveground. The testing at these facilities was crucial to gaining a fundamental understanding of the performance of the sensors but lack the variability of "real sites" and did not focus on detecting leaks belowground on linear assets such as pipelines. Therefore, the overall goal of this project was to advance and validate drone mounted remote sensing technologies designed to detect natural gas leaks in belowground transmission and distribution pipeline systems and identify right of way (ROW) integrity threats via testing at sites representative of the "real-world." With input from the project Technical Advisory Panel (TAP), this overall goal was further focused on leak detection for hard to access real-world sites such as steep slopes and wetlands/water-logged areas, rather than on leak survey for long continuous miles of pipe.

The project focused on a single drone platform offered by SeekOps, Inc. (SeekOps), the project partner, to advance the capabilities of the system, while also developing validation testing and performing thorough/in-depth testing of the system. There are recognized trade-offs, however, in focusing on a single technology, such as not being able to determine how the technology performs in relation to other possible solutions through single or double blind testing as discussed in detail in Ravikumar et al. 2019[7]. Therefore, this work was viewed from the beginning as being a crucial, detailed first step to development and validation of a new, improved technology for use on underground pipeline systems.

Project Goals and Objectives

Multiple targeted goals and objectives were laid out at the outset of this project to accomplish the overall project goal. Table 1 below breaks these objectives down by targeted goals and includes high level steps taken to achieve the goal during the project.

Table 1. Project Goals, Objectives, and Steps Taken for Goal Achievement

Goal	Objective	Steps Taken for Goal Achievement
Goal 1 - Establish a validation testing framework	• **Objective 1a** - Develop a quantitative and qualitative test matrix of technologies. • **Objective 1b** - Determine field test locations.	• Developed test methods that take advantage of Receiver Operating Characteristics (ROC) curves. • Focus of testing was updated to hard to access areas/sites.
Goal 2 - Conduct validation field testing at real-world test sites	• **Objective 2a** - Identify leaks and areas with potentially high consequence integrity threats. • **Objective 2b** - Conduct field measurements and ground truthing. • **Objective 2c** - Collect extensive qualitative and quantitative data on instrument performance.	• Was not able to gain access to high consequence area for testing so adjusted with testing at METEC, one company owned site and one privately owned site. • Conducted three thorough field sampling campaigns to test performance and developed validation framework.
Goal 3 - Obtain a probabilistic understanding of uncertainties for each technology	• **Objective 3a** - Conduct a detailed statistical analysis of the extensive data obtained during testing. • **Objective 3b** – Model uncertainty surrounding key instrument detection field variables (infrastructure, terrain, land cover etc.).	• Conducted detailed statistical analysis of collected data using ROC statistics to test method performance across the three tests. • Impacts of key variables were not modeled due to adaptations of field testing and focusing on hard to access areas.

Drone-Based System Development and Overview

The SeekOps equipment developed and used in this project was crucial to understanding the results from testing the validation framework. The project originally intended to focus on a vertical takeoff and landing (VTOL), fixed wing drone, but as development began, it was decided to move to a rotary unmanned aerial vehicle (UAV, drone). The rotary drone was determined to be most appropriate for the close surveying of pipe sections in different terrain, and thus was chosen instead of the VTOL system, which would fly quickly and at high altitudes. Flying at high altitudes significantly limits the detection capabilities of the drone-based leak detection systems, as mentioned earlier, and FAA regulations prohibiting operation of drones beyond visual line of sight, further limiting the usefulness of the fixed-wing drone. Therefore, the DJI Matrice 300 RTK (M300) was selected as it is the most recent and most advanced rotary UAV intended for commercial inspection and survey use and was ideal for surveying hazardous or difficult-to-access areas of pipeline.

DJI Matrice 300 RTK

The M300 (Figure 1), further, was selected on account of its safety and reliability. Its ability to inspect both open corridors and difficult to reach pipe sections was also a key factor that led to it being chosen for this project. The M300 has a heavy payload capacity as well as 55 minutes of max endurance without a payload and 40-45 minutes with the payload included. An extended range of operation in excess of 15 kilometers (km) coupled with coordinated flight control that allows for multiple operators cemented this UAV as adequately fail-safe. Lastly, the M300 possesses advanced obstacle avoidance technology that allows for a remarkably high level of safety when traversing unknown hard-to-access areas.

Figure 1. DJI Matrice 300 RTK (drone-only) during testing

The M300 had backups for 12 different vital flight components. While a scenario involving hardware failure was unlikely, backups were prepared to come online to ensure safety. Some of the redundant systems are barometers, the inertial measurement units, intelligent batteries, and even the RTK positioning system has dual antennas that can serve to back up the compasses, if needed. The drone also has six pairs of vision sensors to sense changes in position and altitude which can be replaced by infrared time-of-flight sensors if any sensor encounters difficulties. The M300 can perform an emergency landing with just three propellers in the event of a motor malfunction. The drone is equipped with DJI AirSense to further protect any operator by providing real time information about aircraft within 20km of the drone position. Lastly, there is an FPV camera attached to the drone that delivers a live feed at a 145-degree field of view that ensures operators can still have sight if the main payload camera fails.

The one issue encountered with the M300 did not originate with performance, but instead focused on the company DJI. At the time the M300 was selected, there was some discussion in the drone community about the safety of data collected with DJI components. To avoid any potential complications and to secure data collected using the M300, the leak detection and integrity threat systems discussed below were maintained separately and never connected to the DJI software directly.

SeekOps SeekIR™ Methane Sensor

The leak detection/methane sensor that was paired with M300 was the SeekOps SeekIR™ methane sensor, shown in Figure 2. SeekIR™ is a drone-agnostic system originally developed at the NASA/Caltech Jet Propulsion Laboratory. The methane sensor technology fits into a 6-inch-long tube and directly samples the air to detect a methane leak. The sensor is known as a “point sensor,” therefore it and the drone must be flown directly through a plume of leaking methane. This detection method is in contrast to a “path integrated” system detection that only requires the beam of the sensor to pass through plume of leaking methane (rather than the whole system).

The “point sensor” therefore has be flown at lower altitudes as highlighted by the 1 – 3mdetection distance discussed in Ravikumar et al. 2019[7].

Figure 2. SeekIR™ Methane Sensor

The open cavity on the SeekIR™ methane sensor can instantaneously respond to the presence of methane due to high sensitivity and selectivity. The mid-wave 3.2µm infrared (IR) tunable diode laser absorption spectroscopy (TDLAS) technology allows for a sensitivity as low as 50 parts per billion (ppb) generating a consistent 150 ppb minimum detection level. This level of detection allows the sensor to track flow rates above 1 standard cubic feet per hour (scfh) at a distance of 10 meters from an aboveground leak as shown in Ravikumar et al. 2019[7]. The sensor has a weight of 0.5 kg and an internal battery with 6+ hours of run time allowing it to be compatible with and independent of most UAVs. The sensor also contains an internal data logger that provided 900 mHz real time data to develop telemetry of methane levels which was conducted during the period of this project.

Zenmuse H20T Camera

The DJI Zenmuse H20T camera (H20T, Figure 3) was selected for its ability to capture high definition video as well as high quality thermal imagery which could be used in the identification of integrity threats. The camera system contains three individual cameras (sensors) and is equipped with Smart Track auto image following to be helpful in areas where terrain or other obstacles might obscure the line of sight for a pilot.

Figure 3. Zenmuse H20T Camera

To summarize the capabilities of the H20T,

Table 2 provides specifications for each of the three camera types housed in the H20T: zoom, wide angle, and thermal. It should be noted that thermal camera has a spectral range of 8-14µm,

	Zoom Camera	**Wide Angle Camera**	**Thermal Camera**
Sensor	20 MP	12 MP	Uncooled Vox Microbolometer
Lens DFOV	66.6°-4°	82.9°	40.6°
Lens Focal Length	6.83-119.94 mm (equivalent: 31.7-556.2 mm)	4.5 mm (equivalent: 24 mm)	13.5 mm (equivalent: 58 mm)
Lens Aperture	f/2.8-f/11 (normal), f/1.6-f/11 (night scene)	f/2.8	f/1.0
Lens Focus	1 m to ∞ (wide), 8 m to ∞ (telephoto)	1 m to ∞	5 m to ∞
Video Resolution	3840x2160 @ 30fps, 1920x1080 @ 30fps	1920x1080 @ 30fps	640x512 @ 30 Hz
Video Format	MP4	MP4	MP4
Photo Size/Image Resolution	5184x3888	4056x3040	640x512
Image Format	JPEG	JPEG	R-JPEG (16 bit)

with a noise-equivalent temperature difference sensitivity of less than or equal to 50mK at an aperture of f/1.0. This spectral range does not coincide with spectral bands for methane so cannot be used to image methane plumes.

Table 2. Camera Specification for the Zenmuse H20T

Mounting of Systems to Drone

	Zoom Camera	Wide Angle Camera	Thermal Camera
Sensor	20 MP	12 MP	Uncooled Vox Microbolometer
Lens DFOV	66.6°-4°	82.9°	40.6°
Lens Focal Length	6.83-119.94 mm (equivalent: 31.7-556.2 mm)	4.5 mm (equivalent: 24 mm)	13.5 mm (equivalent: 58 mm)
Lens Aperture	f/2.8-f/11 (normal), f/1.6-f/11 (night scene)	f/2.8	f/1.0
Lens Focus	1 m to ∞ (wide), 8 m to ∞ (telephoto)	1 m to ∞	5 m to ∞
Video Resolution	3840x2160 @ 30fps, 1920x1080 @ 30fps	1920x1080 @ 30fps	640x512 @ 30 Hz
Video Format	MP4	MP4	MP4
Photo Size/Image Resolution	5184x3888	4056x3040	640x512
Image Format	JPEG	JPEG	R-JPEG (16 bit)

For the project, SeekOps, Inc. was able to properly attach the SeekIR™ methane detection system and camera with the M300 as shown in Figure 4. The H20T was mounted on the top bracket that was provided by DJI which also allowed the camera to be gimbaled and have free range of motion. Data could be streamed through the drone to the RC ground controller with high resolution data able to be stored on the SD card in the camera. The SeekIR™ methane sensor required a custom mounting bracket that was fit onto the lower M300 mount. This configuration allowed for a high ease of connection when mounting the sensor onto the bracket in addition to a consistent data stream through the 900mHz radio to a SeekIR™ receiver for data logging.

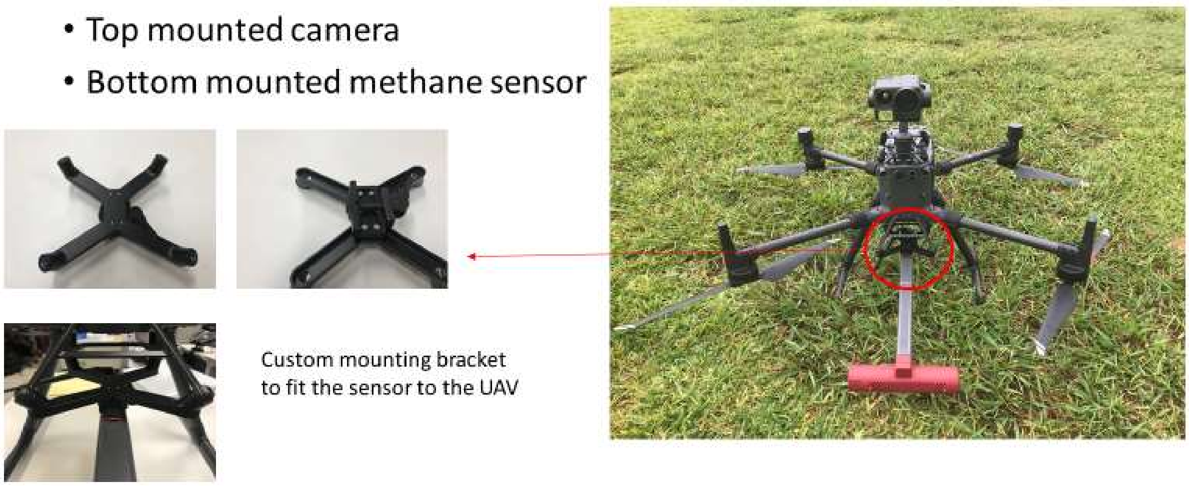


Figure 4. Camera and Methane Sensor Drone Mounting Configuration

Integrity Threat Analysis System

The Vision Lab at the University of Dayton (UD) developed an automated, deep learning-based computer vision algorithm to detect and identify pipeline encroachment threats from aerial images as shown in Figure 5. The automated threat detection designed by UD was combined with the SeekOps image collection drone platform. The UD system was initially developed in partnership with organizations such as Pipeline Research Council International (PRCI) since the early 2010s to develop and fully test the software using images collected by manned aircrafts. Testing at the time of selection had shown that the software had achieved a level of maturity and robustness for analysis of image data collected on drone flights to validate integrity threat identification for basic equipment (vehicles). Future iterations of the software could extend capabilities to cover more configurations of encroachment threat, ground disturbances, and geotechnical threats (outside the scope of this project).

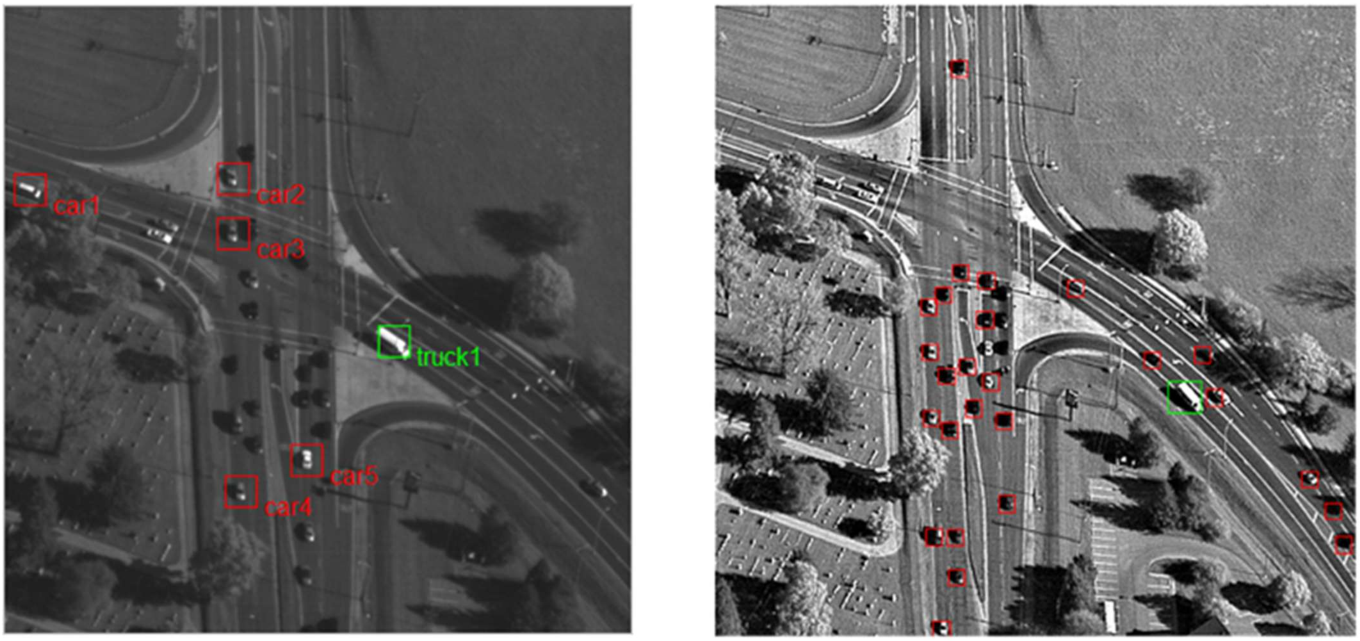


Figure 5. Example detections from University of Dayton software when mounted to a downward facing aircraft mounted camera (photo source University of Dayton Vision Lab).

Technology Evaluation Methodology

The technology evaluation in this project was initially built upon traditional technology validation that was focused on the individual performance abilities of the sensor alone by measuring its accuracy, precision, and time to detect. This sensor-focused approach works well in a laboratory setting with traditional laboratory based analytical techniques but often fails to account for the entire system perspective. The current body of literature for sensors lacks the complex leak detection and integrity threat monitoring systems of sensors that are needed in field scenarios. There are a wide range of systems that could be deployed for leak detection validation: methane detectors, infrared imagers, real-time analytical software, meteorological sensors, and even global positioning systems that can detect and localize potential hazards remotely. Under field conditions the validation of individual sensors is only the first step towards validating full integrity threat and leak detection systems.

Much of the evaluation criteria outlined throughout the course of this project aimed to prove that the full system is validated or that the process of detection is in working order after having already assumed that the individual sensors were tested and validated by a vendor or technology developer. With technology evaluation focused on the methods used rather than individual sensors, developers and users are able to glean much more pertinent information regarding real-world development and deployment. Validation of full systems/methods involves taking a unique look at the evaluation of data from these systems. For this project and to achieve Goal 3, the project team used Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) statistics as a means of evaluating system performance.

Statistical Evaluation/Validation using ROC curves and AUC statistics

ROC curves have been used extensively by the military dating back to World War 2, and more recently have been extensively used in medicine/radiology[9-11]. As explained in Table 3 below, different system indicators like true and false positives and negatives inform the true and false positive rate which allowed the creation of the ROC curve[9-11]. The ROC curve describes the holistic performance of the entire system across various leak detection thresholds. This type of testing will best inform which threshold will bring about the appropriate tradeoff between false negatives (missed leaks) and false positives (false alarms). The area under this ROC Curve (AUC) provides a singular metric to compare results from different systems. Higher AUC values indicate that the system is better at finding leaks while minimizing undesired false alarms – any score greater than 0.5 indicates that the technology being used is superior to a random guess of leak detection.

Table 3. Various term definitions used in the system performance validation

Term	Definition/Description
True Positive (TP)	The system correctly indicates that a leak is present when a leak is present.
True Negative (TN)	The system correctly indicates that no leak is present when no leak is present.

Term	Definition/Description
False Positive (FP)	The system incorrectly indicates that a leak is present when no leak was present.
False Negative (FN)	The system incorrectly indicates that no leak is present when a leak is present.
Leak Margin	The minimum emission rate for a source to be considered a leak. An emission source with a flow rate less than the leak margin is considered a non-leak.
Leak Alarm Threshold	The point (such as a concentration) or scenario (combination of concentration, location, wind direction) above which the instrument indicates there is a leak present. The leak alarm threshold is used in combination with the leak margin. Any concentration above the leak alarm threshold is an indication of a leak that is above the defined leak margin.
True Positive Rate (TPR)	Calculated from the true positives divided by the sum of true positives plus false negatives. Equation - $TPR = \frac{TP}{TP+FN}$
Receiver Operating Characteristic (ROC) curve	A curve created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various leak alarm threshold settings.
Area Under the Curve (AUC)	A summary statistic that represents the area under the ROC curve. AUC is used to reduce the information found in the ROC curve to a single number.

Before delving deeper into ROC curves, it must be noted that in real-world scenarios follow-up foot surveys with handheld sensors would always be required following a remote sensing technology's leak indication. GTI also chose to refine data with different leak alarm thresholds to better understand how balancing false positives and false negatives affects overall measurements.

ROC curves allow for end-users to take a more probabilistic approach in determining which leak margin is best, a question often mistakenly thought to be far more black-and-white. ROC curves provide a way to quantify binary results (detect vs non-detect) across a variable's range (leak thresholds)[12]. ROC curves are satisfactory at visualizing the relationship between true and false positive rates but the AUC statistic provides a quantitative way to summarize a system's performance across a range of detection thresholds. A higher AUC signifies that the system can easily distinguish between true positives and true negatives, the perfect curve would have an AUC of 1.0 (Figure 6, far left).

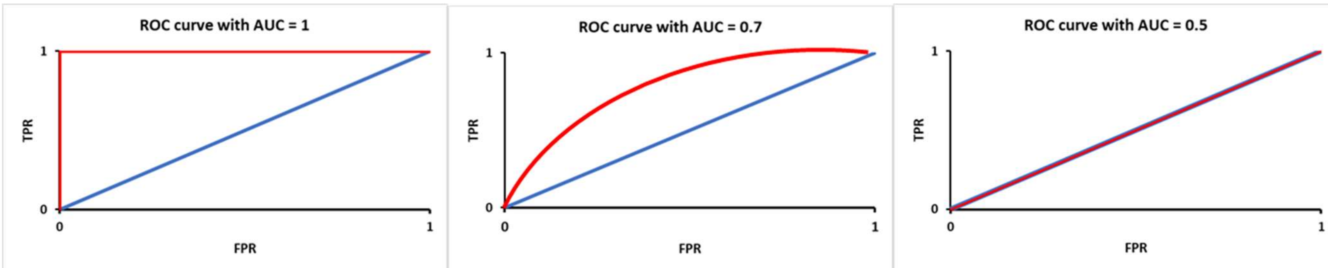


Figure 6. Examples of Types of ROC curves

On the other end of the spectrum, a curve with an AUC of 0.5 (Figure 6, far right) would imply that the system is no better at determining the presence of a leak than a random guess. Typically, distributions will fall somewhere between these two options and settle at an AUC around 0.7 like the ROC curve seen in the middle of Figure 6.

Leak Detection

The validation method for leak detection is predicated on leaks having controllable predefined flow rates that are selected from a "bin" of flow rates. It is equally important that drone pilots do not know the location of different release points so that the tests are "blind." The environment plays an important role –be it wind conditions, precipitation, or near-field obstructions like hedges, trees, or buildings. Many of these obstructions can increase the risk of false negatives but background methane sources near the right of way should also be noted as they increase the risk of false positives. Although these background sources tend to pale in comparison to the controlled release signals during testing, they can overlap in measurements and must be accounted for in the analysis.

Additionally, the number of variables needed to be considered for accurate leak detection had the potential to grow unwieldy when things like meteorological conditions and flight speed are deemed to be variables. To counteract this-flight characteristics and meteorological conditions were treated more qualitatively while the quantitative analysis focused on leak size and terrain type.

Drone System Field Testing Procedure

The design of the testing procedure, and the testing criteria more generally, was meant to clearly outline what factors informed the system for leak detection and integrity threat monitoring validation. The framework for this project was designed for a variety of natural gas transmission pipeline uses in mind so that it can be applied broadly far beyond this project.

Leak Detection Test Procedure/Framework Design

Leak detection is best characterized by how it was carried out within the framework. In this way, Table 4 below contains the necessary information that was recorded to better inform raw and refined data results. As described previously, the leak detection capabilities of the SeekOps integrated platform for transmission pipelines was evaluated as a whole system rather than just focusing on sensor validation. The process of detection is of the utmost importance in this instance because an in-depth analysis of the methods used, and their strengths or shortcomings will provide future end users of the technology with the most important information when considering real-world deployment.

Several of the variables from Table 4 were instrumental in evaluating leak detection performance. The leak margin, importantly, was used to determine what leak flow rate actually was defined as a leak, above which the leak alarm threshold would trigger, indicating the presence of a leak. The leak margin informs the analysis that is done after testing by directly defining what readings qualify as false positives, false negatives, true positives, and true negatives. Therefore, selecting the correct leak margin is key since ROC curves and the AUC statistic run the risk of being skewed if a proper leak margin is not determined.

Much of the data was logged on a one second basis during flight to certify that the selected leak margins were correct and to produce meaningful results. Some of the parameters included in this data collection are locations, altitudes, and battery levels.

Table 4. Summary Table of Variables to be Collected during Drone-Based Leak Detection Validation

Key Performance Indicator (KPI)	Metric	Description
Bin Number	1-6	The bin number represents a grouping of leak rates that will be used during testing. A preliminary suggestion was to use six bins with at least 5 controlled releases falling into each bin.
Leak Margin	Emission Rate (SCFH)	The minimum flow rate for an emission source to be considered a leak.
Leak Threshold	Concentration (PPM) or Scenario	The point (such as a concentration) or scenario (combination of concentration, location, wind direction) above which the instrument indicates there is a leak present.
Terrain Type	Flat, Grass, Rolling, Wetlands, Canyons, etc.	Landscape features that may present unique obstacle and plume migration variation during testing.
Leak Detection	Yes/No	The ability of the instrumentation to detect the presence of a simulated leak above the leak detection threshold.
Detection Max Concentration	Concentration (PPM) or max scenario	The maximum concentration or scenario registered by the instrumentation during detection. This number will be used to determine the probability of performance.

Key Performance Indicator (KPI)	Metric	Description
Leak Survey Start Time	CDT/CST	The time of day (in central time) that the leak survey begins for each controlled gas release point. This variable will (in part) determine the length of time needed to detect the plume.
Leak Identification Time	CDT/CST	The time of day (in central time) that the instrumentation registered a detection at each controlled gas release point. This variable will (in part) determine the length of time needed to detect the plume.
Leak Identification Flight Height	Feet AGL	The height above ground level (AGL) registered by the drone aircraft during detection. Missed leaks may be a function of flight height.
Leak Identification Flight Speed	M/S	The speed in meters per second (M/S) of the drone aircraft during detection. Missed leaks may be a function of flight speed.
Leak Location	GPS Plume and Source Coordinates	The GPS coordinates of the detection point within the plume for the initial scan as well as the GPS coordinates of the source location. If source location requires post processing, the length of time required to localize the leak source is also provided. +/- 5 meters for the plume source will be considered acceptable.
Wind Speed	M/S	Time stamp data for wind speed in mph will be registered from the beginning to the end of the survey for each gas release point. Ability to detect may be a function of wind speed.
Wind Direction	N/S/E/W or Degrees	Time stamp data for wind direction will be registered from the beginning of the survey to the end of the survey for each gas release point. Ability to detect may be a function of wind direction – or plume direction shift due to obstacles.
Atmospheric Stability Class	A-E	The atmospheric stability class will be recorded at the beginning of each 20-minute survey. Ability to detect may be a function of atmospheric stability.
Relative Humidity	Percent	Relative humidity will be recorded at the beginning of each 20-minute survey. Ability to detect may be a function of humidity.

There were slightly different test procedures between the large-scale laboratory testing and the real-world sites which are discussed later. The test procedure and the framework for both kinds of testing are informed by the collection of variables displayed in Table 4. The hypothesis of comparing the two procedures was that the leak detection performance of the system at real-world sites was expected to be less optimal compared to the large-scale laboratory testing, however, as shown below this turned out to not be the case.

Test site selection

Initially, testing was planned to use leak survey of continuous miles of pipelines in urban, suburban, and rural areas which would offer unique perspectives into how the varying complexity of each terrain type might affect detection and localization accuracies. However, after consultation with the TAP and industry partners, there was a shift in focus to testing that may be representative of locating leaks in hard to access areas in general, rather than performing surveys on long continuous miles of pipeline. This shift was spurred on after further discussions with SeekOps and the TAP members who also determined that urban and suburban areas posed far too high of a logistical risk when accounting for current technological capabilities and/or FAA regulations. The TAP members indicated that the most pertinent use cases for drone-based leak survey on transmission pipelines in the near term would be in generally inaccessible areas such as pipelines on steep slopes or in wet and waterlogged areas. In these mountainous regions and wetlands/marshes leak surveying can be difficult and time consuming as they are far less frequented by people which also decreases the likelihood of the leaks to be reported at all. To fulfill Objective 1b (shown in Table 1), Table 5 shows the three test sites that were selected for testing, a brief description of the site, and some challenges unique to that particular site.

Table 5. Test Site Characteristics

Test Site	Site Description	Challenges with Test Site
METEC	Contains real non-operational natural gas equipment with 200+ remote controlled above ground release points. An array of hidden underground leak locations along pipeline structures and ROWs as seen in Figure 7.	Restricted in the number of obstacles and terrain types to test. No real-world practice with large background signals to account for. Potential for highly variable wind conditions in a plane setting.
Utility-owned transmission pipeline	Large field with plenty of navigation room along the pipeline ROW. Near current gas operations equipment causing some background detection.	Background from additional methane sources must be calculated and accounted for. Less precise timings and flow rates of manufactured leaks than in a laboratory setting.
Ranch	Nearby natural gas production and processing equipment, very high background levels when downwind of plumes. Variable terrain with room for a fabricated water source.	Foggy conditions with high moisture in the air tended to carry methane from surrounding sources and saturate equipment. Testing could only be conducted during certain hours. Areas of variable terrain closest to production equipment.

Drone System Field Testing at a Large-Scale Field Laboratory (METEC)

Initial testing was conducted at the Colorado State University (CSU) Methane Emissions Technology Evaluation Center (METEC), shown in Figure 7, a large-scale field laboratory developed under the ARPA-E MONITOR Program. The initial testing was meant to provide foundational information on the ability of the SeekOps drone system to identify underground simulated pipeline leaks in a controlled setting. METEC has approximately 50 belowground leak points installed at 1, 2, and 3 feet below the surface, with 25 points along the pipeline ROWs and the other 25 points in the pipeline test bed.

METEC was chosen to examine a range of leak rates and fully develop an ROC curve and AUC statistics for the SeekOps technology on underground pipeline leaks. Right-of-way integrity threats were also preliminarily investigated at METEC using the UD software. This controlled test allowed for repeatability that can yield statistically significant results. This testing was crucial to obtain foundational information given the potential logistical challenges to testing in difficult/real-world terrain.

Leak detection testing at METEC

Prior to testing at the METEC facility a quantitative and qualitative test matrix for the technologies was developed (Objective 1a). There were two variables focused on for the leak detection experiment: flow rate of simulated leaks and flight altitude of the drone system. Leak flow rate was the primary test variable because it is one of the main factors in leak detection performance. Drone flight altitude was also considered a key factor[5].

Simulated leak flow rates were categorized into six "bins" ranging from 0 to 50 scfh based on discussions with the TAP. One of the leak rate bins was a "blank" for which no gas was released. Categorizing leak rates into bins is done to enable performance comparison across bins and to facilitate the ROC analysis. Higher numbers of bins create better granularity for characterizing performance of the drone system but require a larger sample size in the field test. For the ROC analysis, leak detection was treated as a test with a binary outcome, positive for leak detected or negative for leak not detected.

Table 6. Performance Bins for Methane Leak Detection at METEC

Bin Number	Simulated leak rate (SCFH)	Min. # of releases
1	Blank / 0	5
2	1-10	5
3	10-20	5
4	20-30	5
5	30-40	5
6	40-50	5

Table 7. Flight altitudes focused on for the METEC testing

Description	Flight altitude (m)
Low altitude	2
Mid altitude	5
High altitude	10

Flight altitude was selected as an additional categorical variable in the test since it was considered to have a significant impact on the ability of the drone system to detect underground leaks. Given that the methane sensor on the drone system is a point measurement, the drone had to transect methane plumes in order to find a leak. The lower the drone was to ground level, the higher the methane concentration signal-to-noise ratio captured by the drone when traversing through a plume, and the higher the probability for the drone to find the leak. However, safety is a concern when flying at low altitude. For the METEC testing, three altitudes were selected ranging from 2 to 10 meters. During each round of release, SeekOps was asked to fly starting at low altitude and transitioning to the higher altitudes consecutively.

Prior to the test, SeekOps was asked to share the parameter they monitor to signal a leak detection (e.g., instantaneous concentration or rolling average concentration for last 1 minute) and the threshold value for that parameter (e.g., > 50 ppm). The definition of a leak had to be established as well. For the METEC testing, we defined a leak as any methane emission above 0 scfh. Some stakeholders/pipeline operators may want to change this definition (raise the leak rate definition) in their leak survey strategy to target larger leaks first.

The locations of the release points and the release rates were unknown to SeekOps during each round of releases. Underground leaks were located along the 10 pipeline ROWs at the test facility as shown in Figure 7. Prior to test day, SeekOps was provided the latitude and longitude of the end points of the pipeline corridors (ROWs). This helped SeekOps prepare automated flight plans prior to the test.

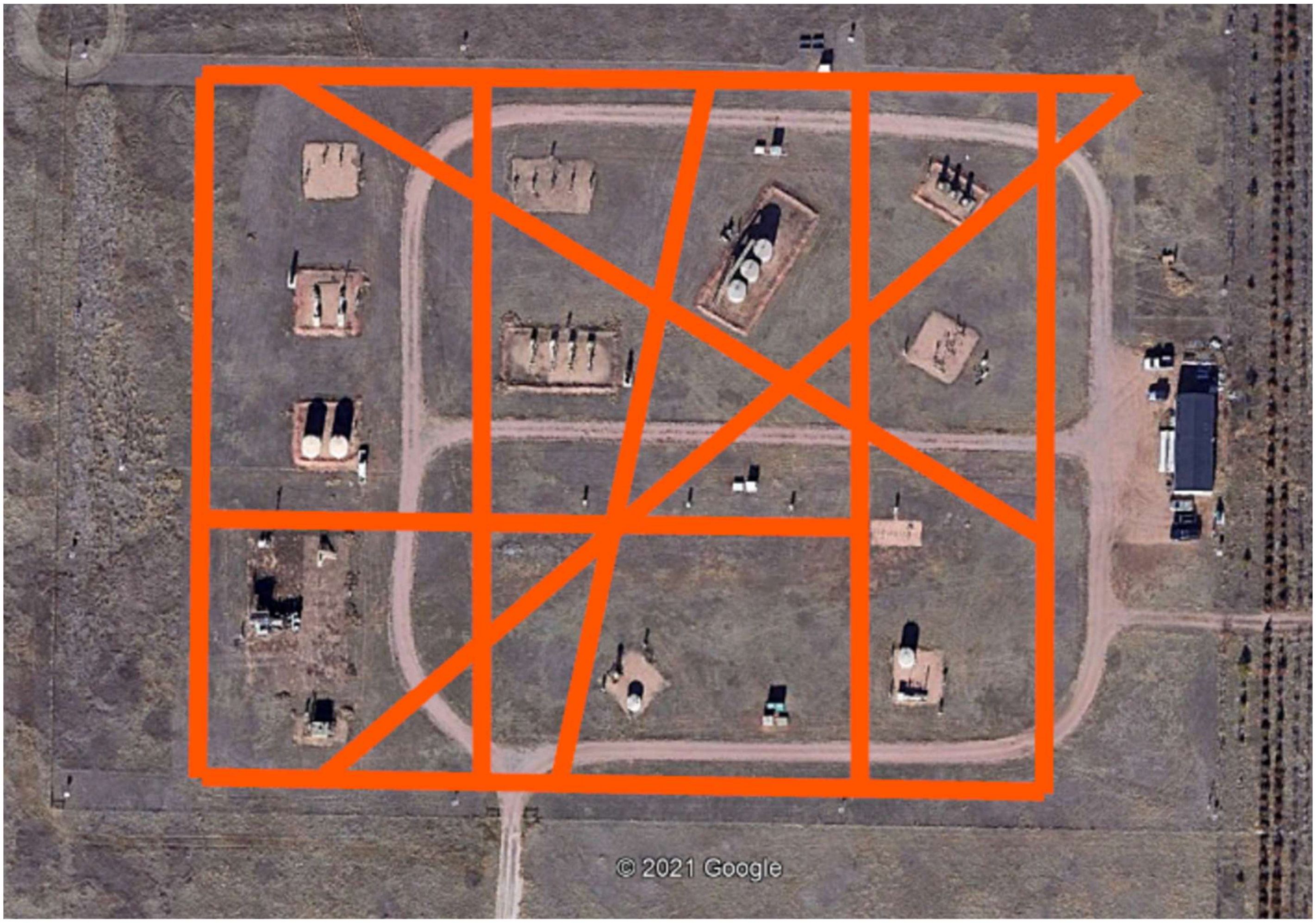


Figure 7. Pipeline right-of-ways at METEC facility at Colorado State University

SeekOps was asked to establish a consistent and automated flight plan to generate results that were comparable. Also, to introduce blanks in the release experiment, there were rounds where no leaks were opened. This approach was useful to evaluate the tendency of the drone system to generate false positive indications or false alarms. Ambient air sampling with a handheld sensor was performed to ensure there was no gas plume along the ROW.

The actual leak detection testing included the collection of data shown in Table 4, and also the survey duration per length of pipeline was monitored to assess "efficiency" of the system and process. This "efficiency" metric served to better understand the likelihood that a given system will find a leak successfully. Flight altitude is important to assess feasibility of survey in hard-to-access areas.

Once data from METEC was collected, the ROC curve for the system and subsequently the AUC statistic was generated. As mentioned earlier, ROC analysis is commonly used to characterize performance of a binary classification task, in this case leak detection or no-leak detection. The AUC statistic describes the performance of the system across the range of flow rates, and across multiple leak thresholds values, into one number. Information from the ROC curve and AUC statistics was used to determine the optimal leak margin and leak alarm threshold to be used in real-world detection.

More specifically, collected data was pre-processed from the raw data form which was a .csv format. Each file contained the time series record of the sensor information including: the timestamp, GPS locations, concentration measurements, altitude, wind speed and wind direction, temperature as well as detection flag. The frequency of data recording was averaged at 10 Hz. The detection flag was calculated using the SeekOps algorithm that determined if the current concentration measurement indicated a methane detection. The most important information was the timestamp, GPS location and the detection flag. Each data file was a continuous recording of several consecutive runs. To conduct the analysis, some pre-processing was needed to parse the data file into runs/passes. It was challenging to extract the individual runs/passes from the raw data. The data was sequentially parsed and categorized into individual runs. After smoothing the raw coordinate data to remove the noise in the GPS signal, the project team:

1. Calculated the heading (traveling direction) of the drone in each time interval.
2. Defined a run when the heading change was greater than 170 degrees within a second.
3. Used the starting and ending point of each run and computed the distance to the end coordinates of the ROW. A run was defined on the ROW that was closest.

Integrity threat monitor testing at METEC

Integrity threat monitoring occurred concurrently with the leak detection testing at the METEC facility where the goal was for the technology to correctly identify encroachment toward a pipeline asset/ROW or identify unauthorized third-party excavation near the site. GTI determined that the optimal way to set this testing up was to place integrity threat identifiers along a defined METEC ROW. Areas along the ROW with vehicles were evaluated to determine information on false positives.

An initial analysis of sample video provided by SeekOps, Inc. from prior work indicated the UD software was capable of identifying vehicle threats even though the drone video was taken from a forward-looking camera. The UD software had been traditionally used on downward looking aircraft mounted cameras. The information from the initial test was used to inform data gathering during the testing at CSU's METEC facility in April 2021. In particular, images of potential threats were collected from both a forward-looking and downward-looking camera to determine methodological differences.

Drone System Field Testing at Simulated Real-World Sites

The simulated real-world site testing framework required a different approach due to the various limitations for a setting where methane releases can only via a controlled release from compressed gas cylinders rather than existing buried natural gas transmission pipelines.

Testing was done at two simulated real-world sites: 1) a natural gas utility owned transmission pipeline site (Figure 8), and 2) at a ranch near other gas processing/production equipment (Figure 9) offering a unique real-world experience involving background interference. Members of the TAP stressed that real-world test sites like these were important because they could incorporate a much wider array of terrains, environmental parameters, and potential interferences when compared to a large-scale laboratory setting.

It was communicated early on in discussions with the TAP that testing on a known leak on a natural gas transmission pipeline was an impossibility as arrangements are immediately made to

Figure 8. Utility owned natural gas transmission pipeline ROW site

Figure 9. Private ranch testing site

fix any identified leaks. There would not be adequate time to deploy all equipment necessary to that area and conduct testing. In addition to this reasoning the project team found that a controlled release with a known flow rate would improve the validation process tremendously. To that end, the preferred method of validation testing was controlled methane releases from an aboveground leak point (covered with a mat) or at a known quantity inside of a water feature.

The procedure for real-world leak detection drone system validation testing at the utility owned natural gas transmission site was as follows:

1. Identify the location of any additional background methane near the testing site and make note of this to better contextualize data analysis.

2. A plan of action was tailored to two different methods of surveying: straight line paths and box methods. Testing was performed at altitudes of 2, 3, and 5 meters with 6 passes.
    a. Straight line paths involved having the drone make straight passes directly above the leak source as well as at a measured distance off-axis of the leak source.
    b. Box methods entailed having a drone fly a square pattern over a determined area for leak surveying.

3. Determine a detection rate/probability for each lake rate bin as shown in Table 6 above and compare results with the large-scale laboratory testing.

At the private ranch test site, we introduced a water test to simulate under water leaks. The test was performed using a box pattern flight.

Integrity threat monitoring testing was only performed at METEC due to restrictions with staging equipment on active/operating sites.

METEC Testing Results

Field testing at METEC was conducted in April 2021. Exact leak locations were not disclosed to the drone operators. Wind conditions were generally mild with an average wind speed of 3.0 m/s throughout the 5-day test period with gusts upward of 5.0 m/s during late afternoons. During the test, 516 passes/runs were completed with 162 runs that had a leak detection and 354 without as shown in Table 8.

Table 8. Test summary and initial statistics at METEC facility

Test Summary		Run/pass Summary	
# of leak rates used	6.0	# of runs/passes	516
# of altitudes used	7.0	# of runs on blank leaks (0 scfh)	45
Average # of runs per scenario	6.8	# of runs on leaks (>0 scfh)	471
Average wind speed (m/s)	3.0	# of runs with detections	162
# of leak sources used	15.0	# of runs without detections	354

A run was considered to have a detection if there was at least one detection above the SeekOps predetermined concentration threshold for a leak in the run. An average of 6.8 runs were performed with each scenario (one combination of leak rate and drone altitude). There were 45 runs performed with "blanks" (no controlled release) and 471 runs with controlled releases. During experimental planning, GTI aimed to get enough passes on blank leaks to determine the false positive performance (chance of creating false alarms). The black dots shown in Figure 10 represent the location records of the drone during all days of testing at METEC.

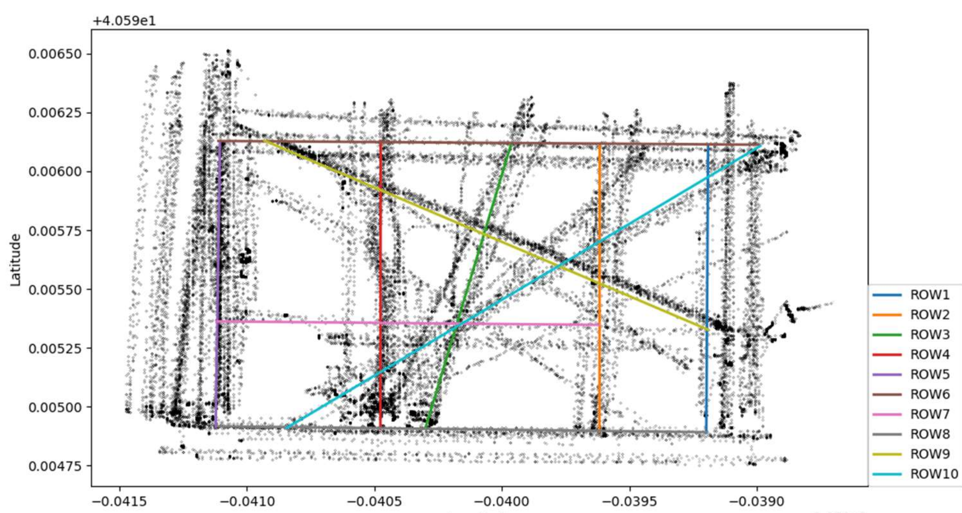


Figure 10. All location records from the drone for the 10 ROWs at METEC

During a “blank” run on the fifth day of the testing, a leak source was found to be emitting. So, the false positives in the blank runs were actually true positives. To eliminate the effect of the unexpected leak, the data from the blank test on the fifth day was removed. To be specific, the data from the 18 runs along ROW 9 on April 30th were removed from the analysis. The updated summary is listed in Table 9. Of the 18 runs removed, nine had a detection and the other 9 did not.

Table 9. Updated run summary after removing contaminated data

Run/pass Summary	
# of runs/passes	498
# of runs on blank leaks (0 scfh)	27
# of runs on leaks (>0 scfh)	471
# of runs with detections	153
# of runs without detections	245

The data analysis was performed based on the updated dataset after removing the contaminated data. The detection probability for each flow rate group was calculated separately. The detection probability was defined as the number of detections divided by the total number of runs in each group. For an actual leak, the detection probability should have been close to 100% (high true positive rate) to be considered a good detection rate, while for a blank leak, the detection probability should be 0 (no false positives).

Runs were grouped to study the potential improvement in detection probability when one or more runs were combined. When grouping, if any of the individual runs resulted in detection then the entire group was considered to have a detection. The reason for grouping passes was that the drone would always go out along the ROW and come back if being used in the real-world. This equals at least two passes along the ROW. It is also advised as best practice to do multiple passes for remote devices. The histograms of detection probability for each flow rate sampled when going a single pass to grouping two and four passes together is shown in Figure 11. There is a trend indicating that the detection probability increased as more runs were grouped. For example, with 4 passes, 45 scfh leaks were detected 100% of the time, while with one pass the 45 scfh leaks were only detected approximately 45% of the time.

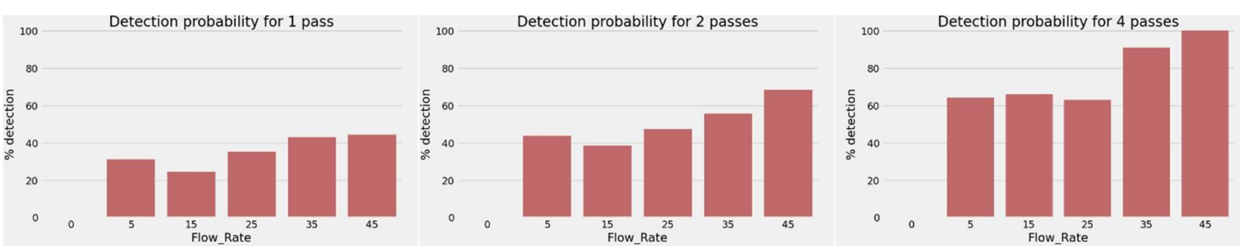


Figure 11. Detection probability for each controlled flow rate for METEC data

Additionally, the leak detection probability was evaluated at different drone altitudes, since as mentioned earlier this could affect the ability of the drone system to detect leaks. The hypothesis was that due to the open nature of the METEC facility (no trees or hills to break the wind), methane plumes would tend to stay close to the ground as they were advected by the wind. Thus, flying at a lower altitude would present a better chance to transect the plume irrespective of wind speed. By comparing the detection probability against different altitudes (in Figure 12, left), it was found that the detection probability was generally higher at 2 and 3 m compared to 5 m. The detection probability as a function of the leak rate for various distances from the pipeline within the ROW is plotted in Figure 12, right. 3 meters from ROW achieved the most consistent detection probability. So, it seemed to be the ideal distance when it is windy based on this result. Further data is needed to draw a clear conclusion, as the wind speed can also contribute to the detectability of the leak.

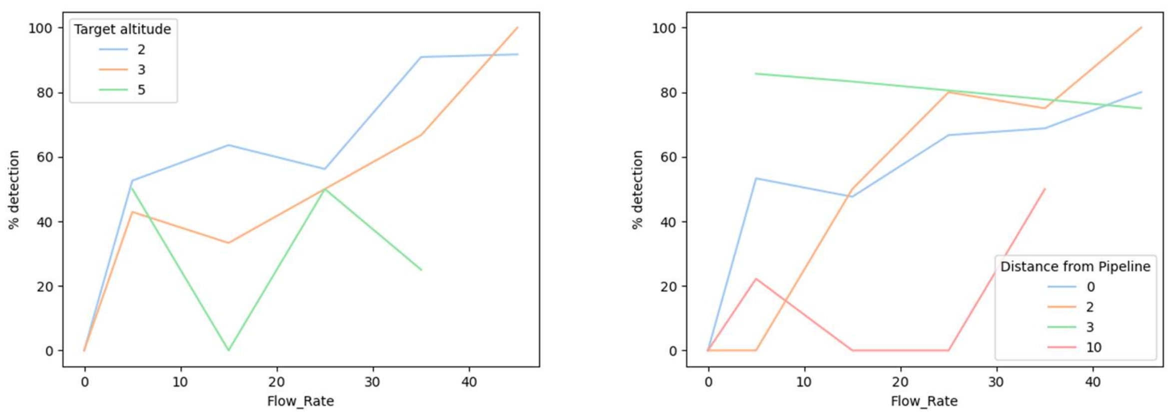
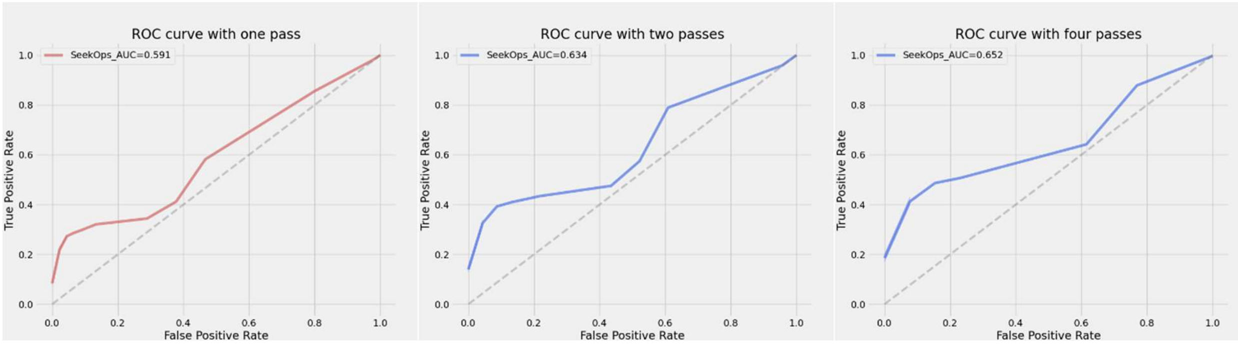


Figure 12. Detection probability as a function of flying altitude (left) and distance from pipeline with the ROW (right) for grouping 4 passes

The results from the METEC testing can be summarized by the ROC curves and AUC statistic. As previously mentioned, the AUC statistic/score summarizes the information contained in the ROC curve and ranges from 0 to 1. Higher values of the AUC score indicate that the system is better at finding leaks while minimizing undesired false alarms. The single pass AUC score for the METEC data is 0.591. The score is greater than 0.5, which indicates that the SeekOps drone technology was better than a random guess at locating leaks at METEC. The AUC score increased as we grouped successive passes indicating that the SeekOps drone technology performance increased when incorporating more passes as shown in Figure 13. The AUC score is listed in Table 10.



(a) (b) (c)

Figure 13. The ROC analysis for METEC data for (a) single passes, (b) grouping 2 passes, and (c) grouping 4 passes

Table 10. AUC score for each group in METEC testing

	Single pass	Two passes	Four passes
AUC score	0.591	0.634	0.652

The relatively low AUC scores highlight several important details of the testing at METEC. The need to 1) sample quickly and 2) for the leaks to remain "blind" lead to two key issues. First the leaks were only allowed to run for a short period of time before the drone was flown over the ROW to be able to maximize the number of sampling runs during the week. Second, the surface expression of the leak was not verified prior to flying the drone. This led to uncertainty in whether the leak was even detectable at the surface, let alone 2, 3, or 5 m in the air, and may have caused some of the poor performance. This was also the first time for SeekOps to be performing this style of leak survey on a belowground asset, so some operating procedures were adjusted during testing, such as where to fly in relation to wind speed and direction to optimally survey the ROW. These factors may have contributed to missed detections during the testing. The evaluation of the performance of the SeekOps drone system technology must consider testing at all of the sites.

Utility Owned Transmission Pipeline Site Testing Results

In June 2021, the teams participated in controlled simulated real-world leak testing at a natural gas utility owned transmission pipeline site. Testing occurred near operating gas transmission equipment (Figure 14) that the team determined to be emitting a detectable amount of methane depending on wind direction and distance from the equipment. Natural gas "sniffers" were used to survey the area to confirm that no other methane sources existed in the test area.

Figure 14. Aboveground equipment at the utility owned transmission ROW site.

Building on the information learned at METEC, two drone flight methods were employed during the simulated real-world site testing: straight line paths and box patterned methods. In straight line paths the drone made straight passes directly along the ROW. For box patterned methods the drone flew a rectangular pattern over an area to survey for a leak, simulating flying down one side of the ROW for set distance, crossing the ROW and flying back to the operator down the other side of the ROW. This procedure could optimize detection based on wind. Both methods were performed at altitudes of 2, 3, and 5 meters and 6 passes were made for each method.
Due to the simplified layout at the utility owned transmission site, the data was less complicated than the METEC data. To enable data analysis, the raw data still needed to be parsed into individual runs/passes. Figure 15 shows the controlled leak location used, along with collected data points. The line test is shown on the left of Figure 15 and box test on the right. This time, due to the simplicity, we used latitude to classify runs. The line test was one traversal along the ROW. The break point for each run was determined to be the change of sign for the slope in the latitude vs. time plot. In the box test, the drone traveled in a squared pattern. The rectangular route consisted of two pairs of horizontal edges and vertical edges (in a longitude-latitude plot).

The change of sign in the slope of the latitude vs. time plot to determine the break point for the box pattern as well. One box test consisted of two sections between the break points.

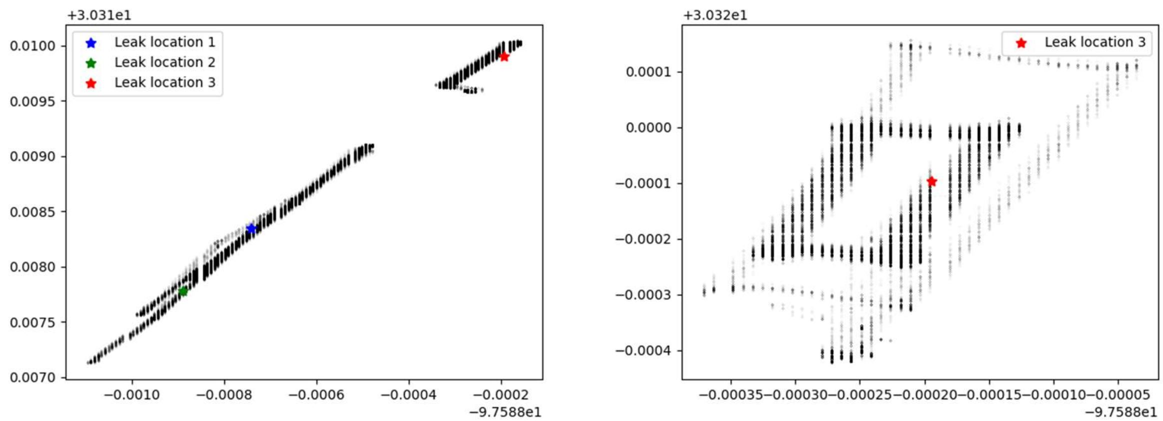


Figure 15. The location records from drone and the controlled leak location for line test (left) and box test (right)

Summary data for the runs can be seen in Table 11 and

Table 12. A total of 390 runs were performed for the line test with an average of 21 runs for each scenario (one combination of leak rate and drone altitude). There were 93 runs performed for "blanks" (no controlled release) and 294 runs on controlled releases. The box tests included 214 runs with an average of 11.9 runs for each scenario with 34 runs on "blanks" and 180 on controlled releases. During experimental planning, GTI aimed to get enough passes on blank leaks to monitor the false positive performance (chance of creating false alarms). A positive run was defined as at least one detection flag from the SeekOps algorithm in a run.

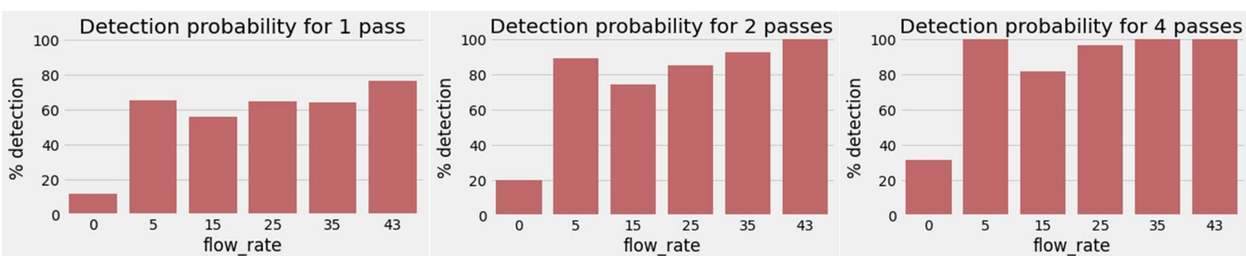
Table 11. Test summary and initial statistics for line tests at a utility owned transmission site

Test Summary		**Run/pass Summary**	
# of leak rates used	6.0	# of runs/passes	390
# of altitudes used	3.0	# of runs on blank leaks (0 scfh)	93
Average # of runs per scenario	21.0	# of runs on leaks (>0 scfh)	297
Average wind speed (m/s)	1.6	# of runs with detections	232
		# of runs without detections	158

Table 12. Test summary and initial statistics for box tests at a utility owned transmission site

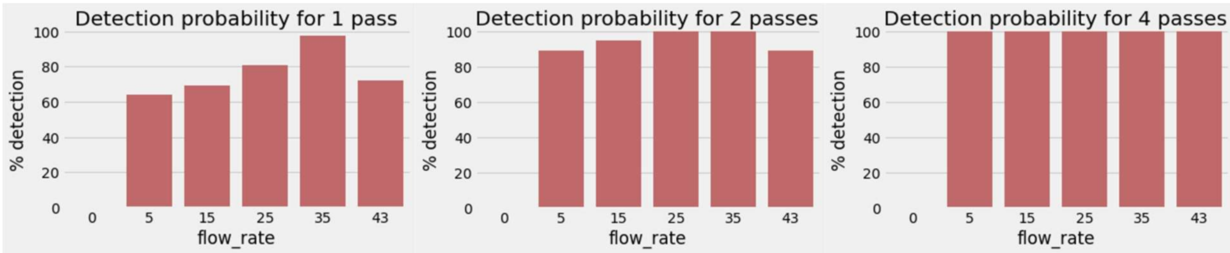
Test Summary		Run/pass Summary	
# of leak rates used	6.0	# of runs/passes	214
# of altitudes used	3.0	# of runs on blank leaks (0 scfh)	34
Average # of runs per scenario	11.9	# of runs on leaks (>0 scfh)	180
Average wind speed (m/s)	0.9	# of runs with detections	159
		# of runs without detections	55

The detection probability for each flow rate was calculated separately. When passes were grouped, if any of the individual runs resulted in a detection, the entire group would have a positive detection. The histograms in Figure 16 and Figure 17 show how the detection probability increased by leak flow rate as more passes were considered. By grouping 4 passes, the box test achieved 100% detection probability and the line test also achieved 100% detection at 5, 25, 35 and 43 scfh leak flow rates. The false positives (detections in the zero scfh rate bin) from the line tests were driven by emissions from aboveground equipment (a metering and regulating station) located at the test site. This provided an opportunity to test in real world conditions, where leaks from assets not being surveyed may complicate leak detection. There were no false positives for the box tests, which could indicate that either this methodology is more effective at eliminating false positives, or that the wind had shifted during the box testing, eliminating the methane plumes from the test area.



(a) (b) (c)

Figure 16. Detection Probability charts for line tests by (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes at the natural gas transmission pipeline site



(a) (b) (c)

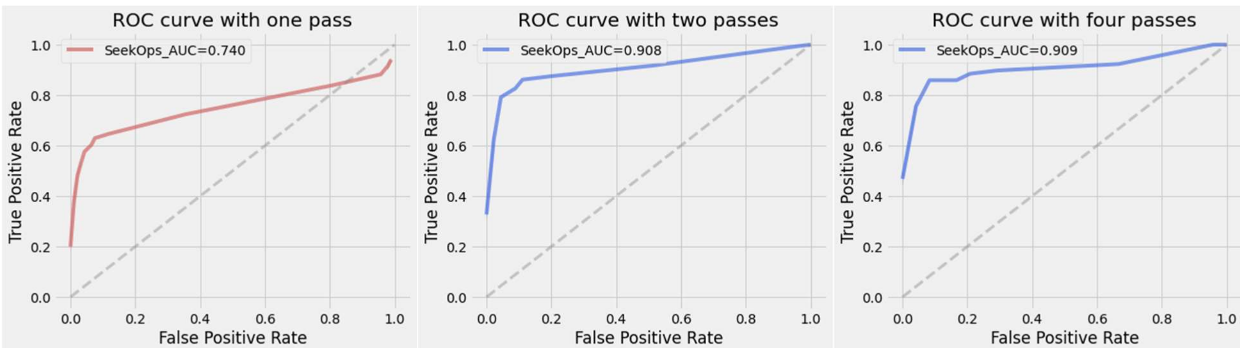
Figure 17. Detection Probability charts for box tests by (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes at the natural gas transmission pipeline site

Additionally, the leak detection performance was evaluated at different drone altitudes. From previous results from the test at the METEC facility, we found the detection probability was higher when the drone was flying at 2 m and 3 m above ground (Figure 12, left). Hence, we concluded that flying at a low altitude presents the best chance to transect the plume. However, that was not necessarily the case in this testing. By comparing the detection probability against different altitudes, we found no obvious benefit of flying at a lower altitude for the line tests as shown in Figure 18. This was likely due to the testing environment. The surroundings at this test site included a lot of trees and bushes. The gas was likely swirling around and contained in the vicinity, which may have made detection easier. While the METEC facility is in an open field, the gas plume is affected more by the wind. Additionally, the box test achieved perfect performance at every height flown, so there was not enough evidence to determine differences by flying altitude. This indicated that the effect of altitude on the detection probability may be related to the type of environment of the test. The METEC facility was in an open field while this real-world facility had more bushes and trees in the surrounding environment.

(a) (b)

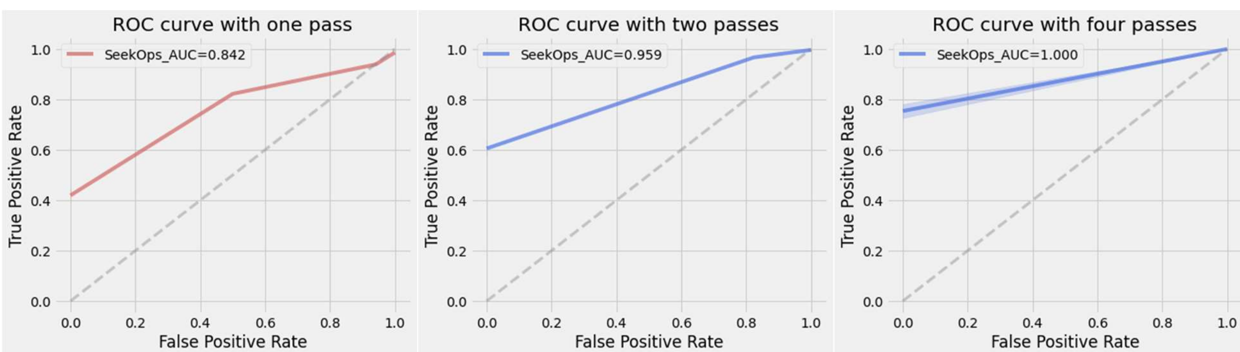
Figure 18. Detection Probability as a function of drone altitude for 4 passes for (a) line tests and (b) box tests

An ROC analysis was completed on the real-world data for both line tests and box tests and is summarized/represented by the AUC statistic. The results for both the line test and the box test are plotted in Figure 19 and Figure 20. Compared to the results from the METEC data shown earlier, the SeekOps drone technology performed much better in this round of tests. The goal for any technology is an AUC score of 1, which means all leaks were properly identified (high true positive rate) with no false positives. In the line test, the AUC score for single passes was 0.740, which is an improvement over the METEC data. Table 13 summarizes the AUC score for different test types and different groups of runs. It can also be seen that the AUC score increased as the runs were grouped together. This agrees with previous findings that grouping runs can increase the detection probability. The AUC score for the box test is higher than the line test and can achieve a perfect score when grouping 4 runs. This agrees with previous findings that the box pattern performs better than the line pattern.



(a) (b) (c)

Figure 19. The ROC analysis for line test at the utility owned transmission site for (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes



(a) (b) (c)

Figure 20. The ROC analysis for box test at the utility owned transmission site for (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes

Table 13. The comparison of AUC score for different tests and group of passes

Test type	Single pass	Two passes	Four passes
Line test	0.740	0.908	0.909
Box test	0.842	0.959	1

The box pattern method achieved better performance than the normal straight-line path. One possible explanation is that the box test involves a circular motion around the leak location, which makes it able to detect any plumes that may be advected by wind. In many cases, the drone detected a leak when it was downwind of the leak release point as shown in Figure 21 where the purple circles indicate the flight path of the drone on the run and the arrow originates from the location of the release and the arrow is pointing along the wind direction heading (in this case northwest).

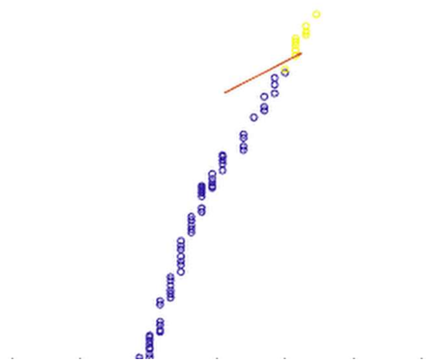


Figure 21. Drone flight path indicating leak detection (yellow circles) when the drone was downwind of the release point.

There were, however, no clear correlations between wind direction and leak detection. In fact, for some runs the drone would register a detection when it was upwind of the release point as shown in Figure 22, where the yellow circles indicate the position of the drone when it detected the leak with the orange arrow originating from the release location indicating the wind direction.

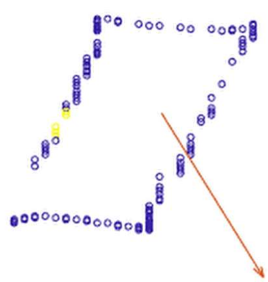


Figure 22. Drone flight path indicating leak detection (yellow circles) when the drone was upwind of the release point.

The lack of relationship between detection location and wind direction highlights an important caveat to leak detection and localization with any drone system. The localization is only as good as the data that is fed to the algorithm. For example, as shown in Figure 8 and Figure 14 above, the area around the ROW was lined with trees that were taller than the altitude of the drone flight paths. This could have caused the local micrometeorology to be variable across short distances, especially near the canopy edge. Further, the wind speed and direction are not measured directly on the drone, but at a fixed anemometer adjacent to the testing area as shown in Figure 23. These factors are likely the cause of the lack of relationship between detection and wind direction.

Figure 23. Utility owned transmission test site with wind monitoring equipment highlighted in the red box.

Real-world testing at a private ranch

In September 2021, the teams visited a second simulated real-world test site on a private ranch. The private ranch field site allowed for the collection of data on a transmission ROW that was near other gas processing/production equipment as shown earlier in Figure 9 and in Figure 24 below. The terrain was also hilly. Emphasis of the test centered on surveying over hilly terrain and over a water source (Figure 24). As shown in Figure 24, concentrations in the background air were as high as 20 ppm or more as measured with the drone and a handheld combustible gas indicator (CGI). This allowed for testing and detection evaluation of the methane sensor in a high background concentration environment.

Figure 24. Private ranch testing near oil and gas production equipment and on the edge of a pond.

For surveying over a water source, a container was filled with water from a pond and tubing was placed in the container and weighed down to keep it under the water surface as shown in Figure 25. Due to landowner restrictions the release point could not be in the pond itself which greatly limited the depth of the release. The drone flew a box pattern around the water source to evaluate detection capabilities. The box pattern tests were performed as in the previous field test at altitudes of 2,3, and 5 meters with 6 passes or repetitions.

Figure 25. Underwater emission release point at the private ranch

One interesting observation was that dense fog and high moisture (as shown in Figure 26) caused high concentrations of methane to be measured in the air (above 50ppm). The ambient air concentrations were above the maximum detection limits of the highly sensitive SeekOps SeekIR™ sensor. The concentrations measured in air with a CGI reached 60 to 80 ppm and they lasted for approximately 2 hours before they dropped below 40 ppm (as measured with the CGI). It is highly likely that a stable boundary layer/capping inversion occurred overnight trapping methane being emitted in the surrounding area, impacting the measurements. This could indicate that in an area of high oil and gas activity that measurements must occur after mid-morning when the nocturnal boundary layer has been replaced due to surface heating and turbulent mixing[13].

Figure 26. Dense fog and high moisture on one day during the private ranch testing.

Overall, there were three types of tests performed at the private ranch - the line test, the box test, and the box test with water feature. Two different controlled leak locations were used for the box test, one location was used for the line test, and one was used for the water feature test. Figure 27 shows the flight patterns around each leak location used for each test type.

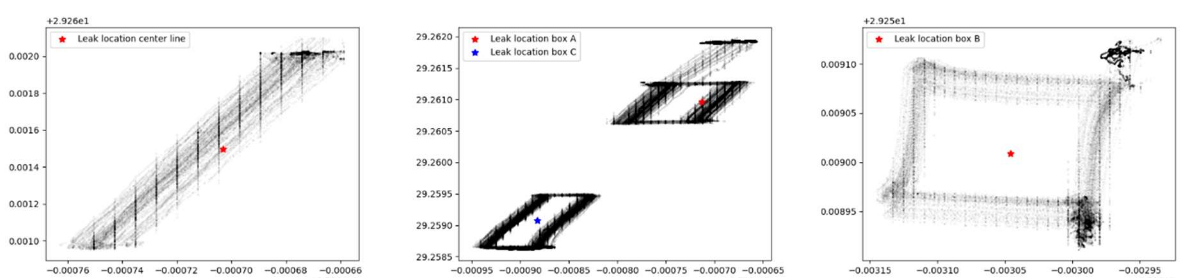


Figure 27. The location records from drone and the controlled leak location for line test (left) and box test (middle) and box test with water feature (right).

The raw data was parsed into individual runs for analysis like the utility transmission pipeline site data with summary statistics given in Table 14 through Table 16. For line tests, there were

114 runs giving an average of 6 runs per scenarios (one combination of leak rate and drone altitude) with 24 of the 114 runs performed on "blank" (no controlled release) and 90 runs on controlled releases. The box tests had 216 runs giving an average of 12 runs for each scenario with 36 performed on "blank" and 180 runs on controlled releases. The box test with water feature had 108 runs giving an average of 6 runs for each scenario with 18 performed on "blank" and 90 runs on controlled releases. The wind was light with an average of 2, 1.5 and 0.9 m/s for the three test types, respectively.

Table 14. Test summary and initial statistics for line tests at a private ranch

Test Summary		Run/pass Summary	
# of leak rates used	6.0	# of runs/passes	114
# of altitudes used	3.0	# of runs on blank leaks (0 scfh)	24
Average # of runs per scenario	6.3	# of runs on leaks (>0 scfh)	90
Average wind speed (m/s)	2.0	# of runs with detections	69
		# of runs without detections	45

Table 15. Test summary and initial statistics for box tests at a private ranch

Test Summary		Run/pass Summary	
# of leak rates used	6.0	# of runs/passes	216
# of altitudes used	3.0	# of runs on blank leaks (0 scfh)	36
Average # of runs per scenario	12	# of runs on leaks (>0 scfh)	180
Average wind speed (m/s)	1.5	# of runs with detections	178
		# of runs without detections	38

Table 16. Test summary and initial statistics for box tests with water feature at a private ranch

Test Summary		Run/pass Summary	
# of leak rates used	6.0	# of runs/passes	108
# of altitudes used	3.0	# of runs on blank leaks (0 scfh)	18
Average # of runs per scenario	6.0	# of runs on leaks (>0 scfh)	90
Average wind speed (m/s)	0.9	# of runs with detections	64
		# of runs without detections	44

As with the other tests, the detection probability for each flow rate was calculated for the three different test types with runs grouped to explore changes to the detection probability and is shown in Figure 28, Figure 29 and Figure 30. The SeekOps drone-based system achieved a

100% probability of detection when grouping four runs in all three types of testing. However, false positives (detections in the 0 rate bin) were detected during all three types of testing. Similar to the findings at the utility transmission pipeline test, the box pattern achieved the best performance even with single passes.

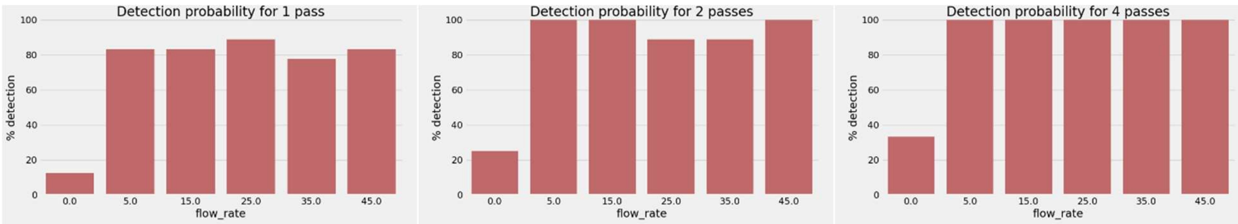


Figure 28. Detection Probability charts for line tests at the private ranch by (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes

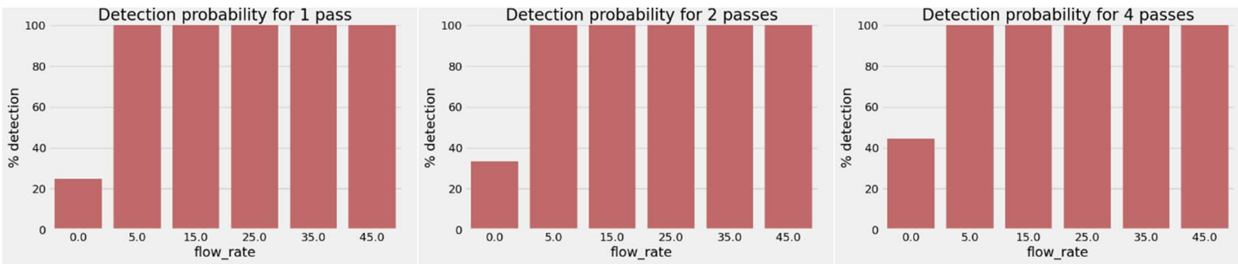


Figure 29. Detection Probability charts for box tests at the private ranch by (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes

The testing conducted with the methane release point under water did appear to have a small impact on the detection probability when using only one pass, with lower detection probabilities than for the box tests that were above water as shown in Figure 30. These differences will be further explored in the ROC curve analysis.

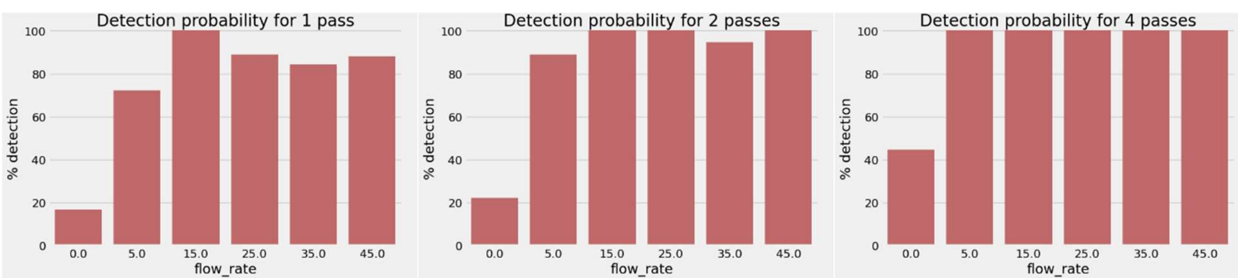


Figure 30. Detection Probability charts for box tests with water feature at the private ranch by (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes

More false positives were detected during this testing than the other two sites. Out of the 36 runs on blank leaks for the box testing (not below water), there are 7 runs with detection flags. The 7 false positives all came on one day of testing with 5 of the runs containing the false positives shown in Figure 31. The dots in the figure shows the drone location with the yellow shaded dots indicating detections and purple dots are non-detects. We can see that the detections are all around the same area to the south of the leak location. Detections that are that consistent would likely indicate an existing leak to the south of the testing location that might be affecting the test results. This indicates that the rigid testing protocols to precisely detail the performance of the drone system, may not be entirely consistent with the actual procedures that would be used in the field. A detection that strong would likely lead to further investigation, such as flying the drone along an adjacent box to the south to determine whether a leak is present in that area. The ability to follow detections in this is an important advantage of drone detection.

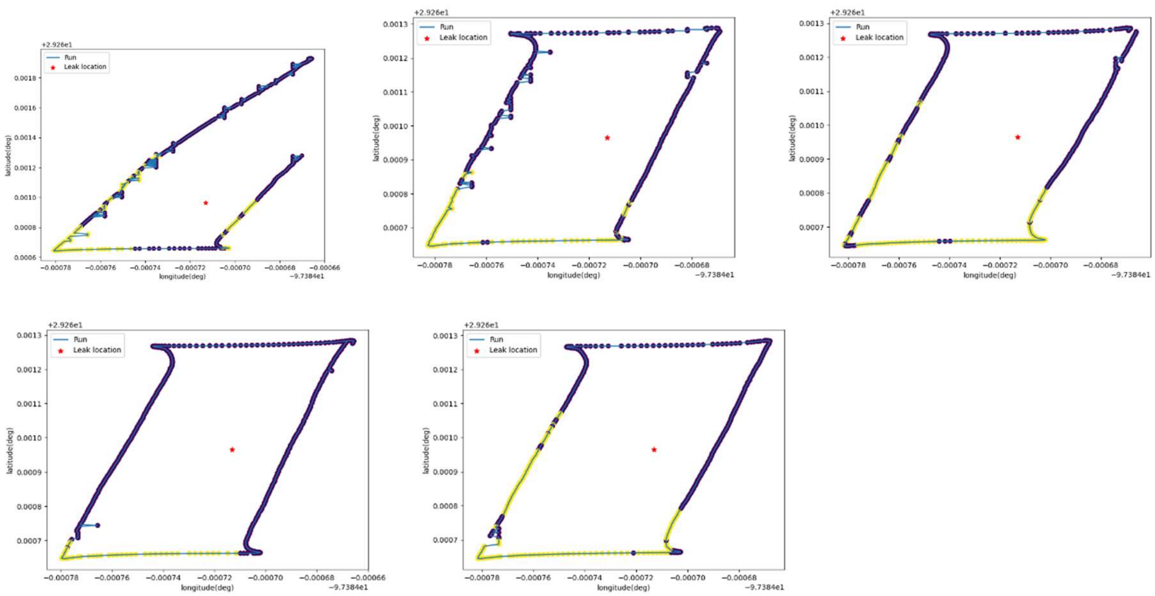
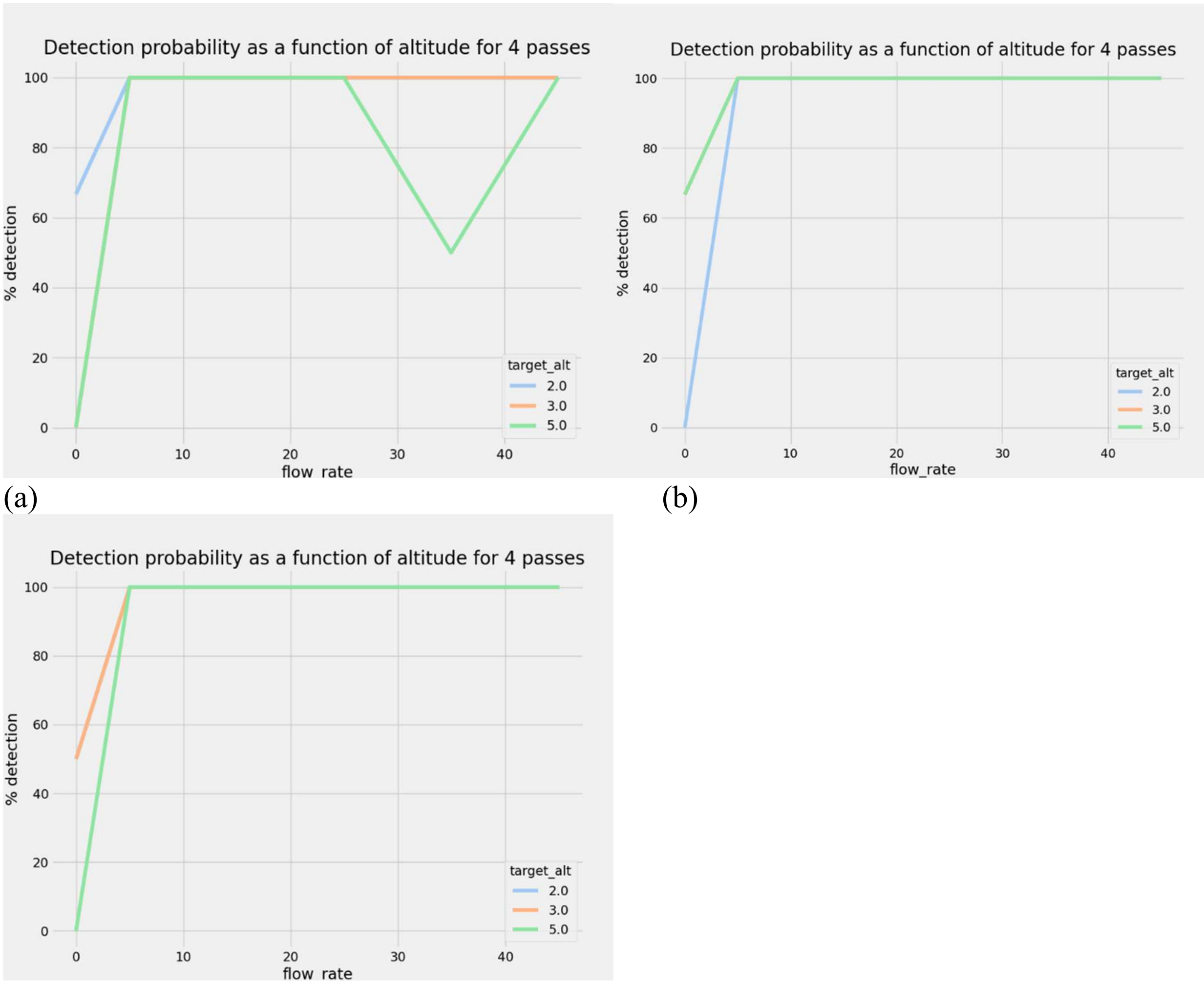


Figure 31. False positives located at south side of leak location for box test on Sep 27th, 2021

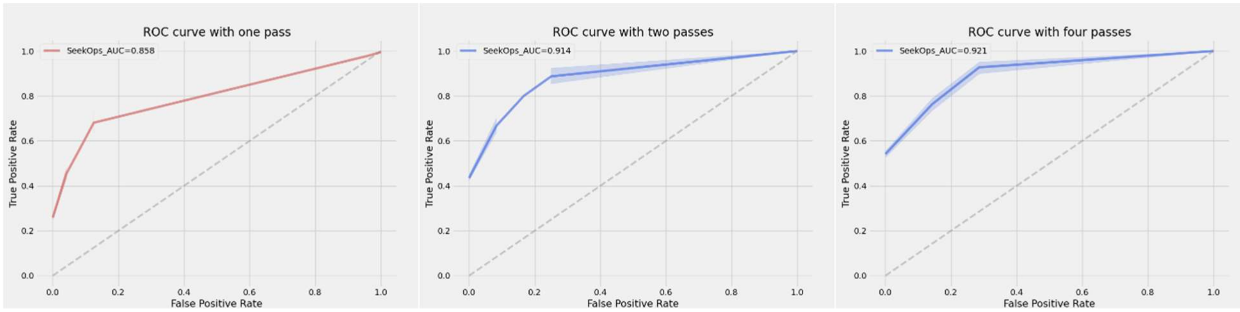
Figure 32 showed the detection probability as a function of altitude when grouping 4 passes in each testing type. Comparing the detection probability against different altitudes, it was found that there is no obvious benefit of flying at a lower altitude for the line tests. The leak detection performance was evaluated at different drone altitudes. Similar to the findings in the transmission site testing, there is not enough evidence of an optimum flying altitude.



(a) (b) (c)

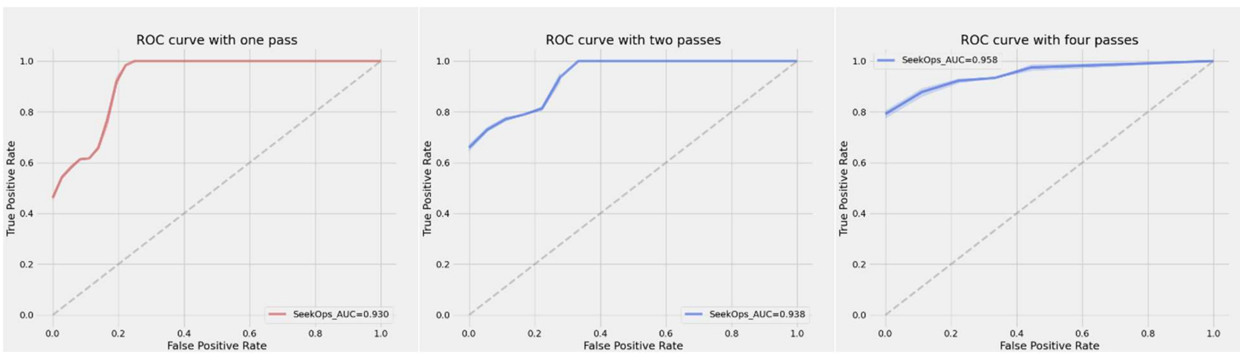
Figure 32. Detection Probability as a function of drone altitude for grouping 4 passes for (a) line tests, (b) box tests and (c) box tests with water feature.

ROC analysis was also performed for the three testing types at the private ranch and summarized with an AUC statistic. The results are plotted in Figure 33, Figure 34 and Figure 35 for the line tests, and box tests with water feature, respectively. The results were quite promising for the SeekOps drone system for the overall testing at the private ranch. As shown in Figure 17, the lowest/worst score was 0.858 for the single pass AUC for the line test, which already indicated good performance before combining runs. When runs were combined, the AUC score improved, even reaching 1, for the testing conducted under water. This indicated that the water had little impact on the performance of the SeekOps.



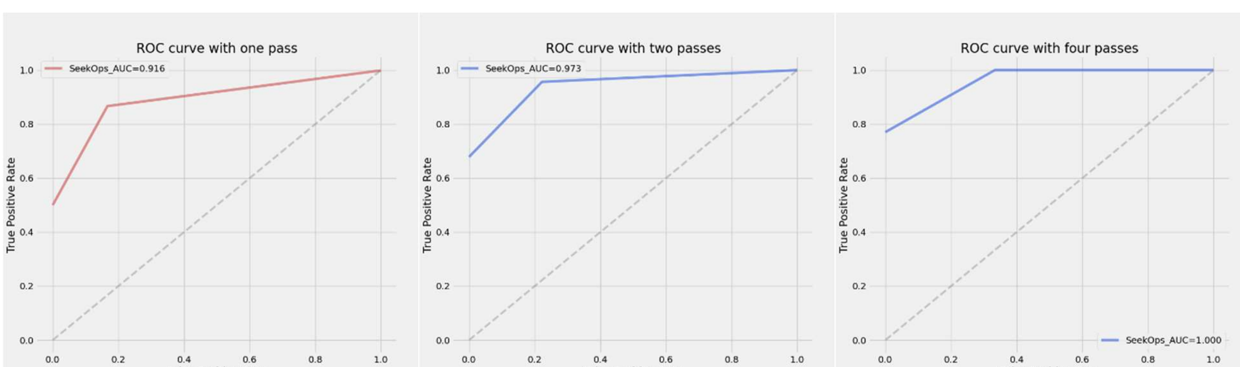
(a) (b) (c)

Figure 33.The ROC analysis for the line tests at the private ranch for (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes for the private ranch testing



(a) (b) (c)

Figure 34.The ROC analysis for the box tests at the private ranch for (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes for the private ranch testing



(a) (b) (c)

Figure 35. The ROC analysis for the box tests with water feature at the private ranch for (a) single passes, (b) grouping 2 passes and (c) grouping 4 passes for the private ranch testing

Table 17. The comparison of AUC score for different tests and group of passes

Test type	Single pass	Two passes	Four passes
Line tests	0.858	0.914	0.921
Box tests	0.930	0.938	0.958
Box tests with water feature	0.916	0.973	1

Although there appeared to be little impact on the Seekops performance when the leak was below the surface of the water, there were still some unknowns to address with this testing. In particular, the leak was under less than 0.5 m of water, which may not be indicative of the depth of transmission pipelines under water. Leaks at deeper depths may be harder to find, especially small leaks, as very small amounts of the methane may dissolve in the water. More work is needed on the impact of underwater leaks on the performance of the drone to be conclusive.

Integrity Threat Monitoring Evaluation at METEC

The UD Vision Lab software was under development, therefore, the libraries of various types of equipment had not been completed for both downward and forward-looking videos at the time of the evaluation. Despite this limitation, the initial evaluation of the software was successful in demonstrating the capabilities, limitations, and usefulness of the software for integrity threat monitoring. Since the system was still under development a high-level qualitative evaluation was performed for video data collected during testing at the METEC facility as shown in Figure 36 through Figure 41.

Figure 36. The METEC testing facility on a blank run with no integrity threats staged

Figure 37. METEC facility with a single pickup truck staged along the ROW.

Figure 38. UD software identification of potential threats (passenger vehicle) near the office building as the drone approaches.

Figure 39. Demonstration that orientation impacted the ability of the UD software to identify threats.

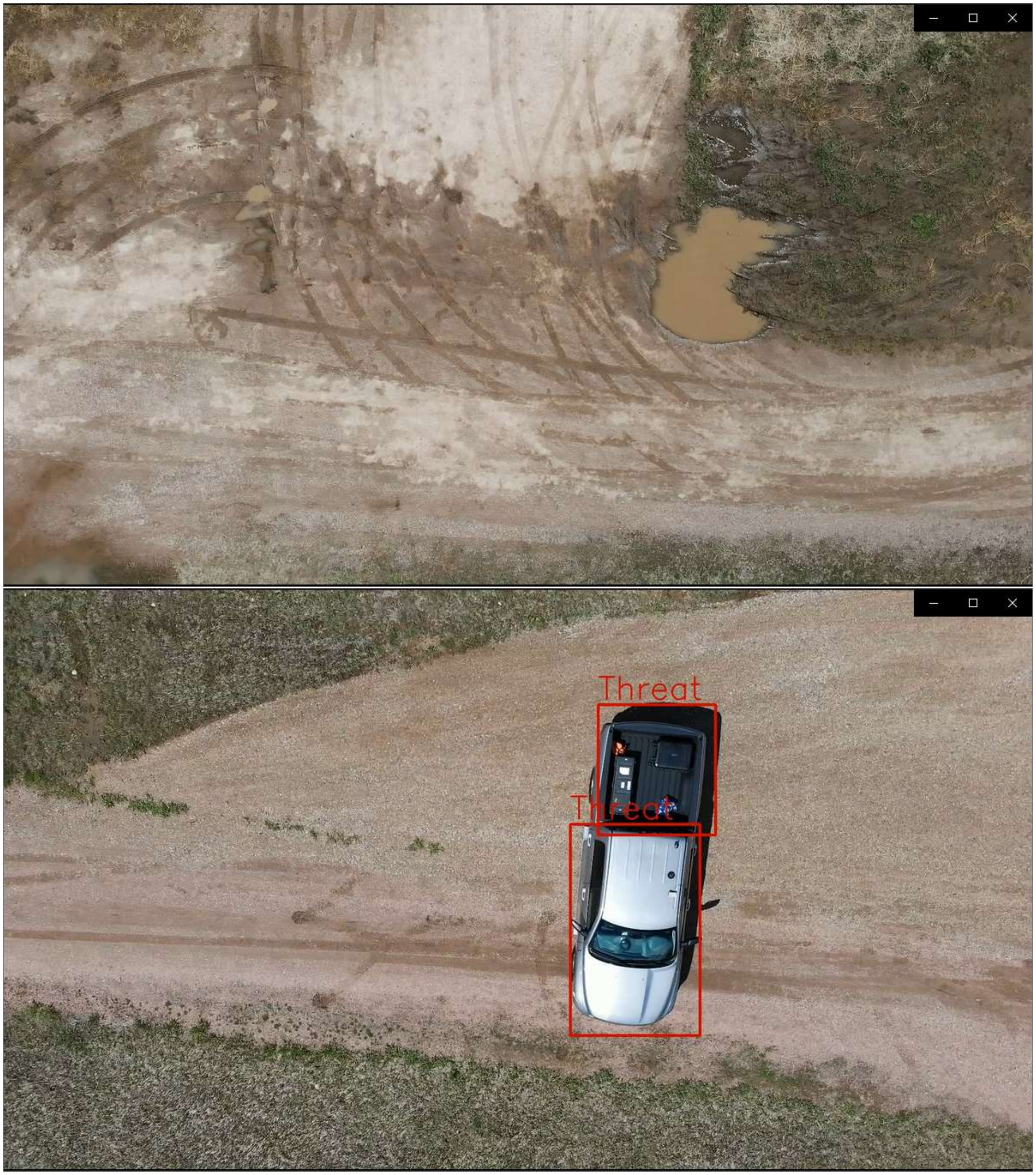


Figure 40. Example of threat detection with the UD software looking directly downward from the drone.

Figure 41. Another example of the UD software identifying a threat. In this case the threat was moving along the road.

The software showed promise for identifying passenger vehicles in the ROW as potential integrity threats. However, more development of the software would need to occur to build the image libraries for a whole host of example integrity threats before a full evaluation could be completed.

Conclusions

Overall, the main goal of the project – to advance and validate drone mounted remote sensing technologies at sites representative of the “real-world” – was accomplished via development of 1) the SeekOps drone system and 2) whole system evaluation methods. We demonstrated that the ROC curves and AUC statistics were useful for comparing the performance of the system across various test scenarios making them an effective method to “grade” or evaluate system performance.

The SeekOps drone technology showed promise for detecting belowground natural gas transmission pipeline leaks across the range of leak rates studied (0 – 45 scfh). According to our testing, we learned that for the drone to be most effective (i.e., have the highest AUC scores across test scenarios) it should 1) make multiple passes across the area being surveyed and 2) fly in a box pattern around the around being surveyed. When doing these two things the drone system could achieve in some cases AUC scores near 1, indicating a perfect tradeoff between true positives and false positives.

In addition, the UD integrity threat monitoring software showed promise for being able to locate potential integrity threats among a crowded landscape at METEC. This indicates that given the time to allow the system to train on various images of potential integrity threats (e.g., heavy equipment, field trucks, people) the system could progress into something quite useful.

The evaluation testing did identify interesting issues that must be considered when performing evaluation testing and using remote sensing techniques to locate leaks. First, the presence of fog or an atmospheric capping inversion can cause local atmospheric methane concentrations to increase to levels that may make detection impossible at times when in areas of heavy oil and gas activity. Second, rigid survey procedures can limit the ability of the drone system to properly locate leaks, however, structured surveys allow for the proper coverage of all areas to be surveyed. In the real world, to take advantage of the ability of the drone to cover large areas there should be a combination of first flying a predefined path to ensure proper coverage, then targeted investigations with the drone in areas where there were detections. This further highlights the need for a rotary drone for the “hard to access” scenarios studied in this project.

One area that was not able to be fully examined in this evaluation study was the ability of the drone system to detect leaks that occur on piping submerged in water, including whether the presence of water can impact the way the leak appears at the surface of the water or whether other factors may impact the ability of the drone to detect a submerged leak. Future work should focus on a full evaluation of various leak sizes at different depths below water to explore any impacts that submerged infrastructure may have on the ability of the drone to detect leaks. Also, further development and testing of the drone mounted integrity threat monitoring system is needed.

References

1. ITRC *Evaluation of Methane Detection Technologies* (*https://methane-1.itrcweb.org/*); 2018.

2. PHMSA Small Scale DIAL for Methane Detection. Contract #DTPH5615HCAP11L. https://primis.phmsa.dot.gov/matrix/PrjHome.rdm?prj=630.

3. DOE APRA-E Methane Leak Detection System. https://arpa-e.energy.gov/?q=slick-sheet-project/methane-leak-detection-system-0.

4. Stanford Mobile Monitoring Challenge. https://methane.stanford.edu/.

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

6. Fox, T. A.; Barchyn, T. E.; Risk, D.; Ravikumar, A. P.; Hugenholtz, C. H., A review of close-range and screening technologies for mitigating fugitive methane emissions in upstream oil and gas. *Environmental Research Letters* **2019,** *14*, (5), 053002.

7. Ravikumar, A. P.; Sreedhara, S.; Wang, J.; Englander, J.; Roda-Stuart, D.; Bell, C. S.; Zimmerle, D.; Lyon, D. R.; Mogstad, I.; Ratner, B.; Brandt, A. R., Single-blind inter-comparison of methane detection technologies – results from the Stanford/EDF Mobile Monitoring Challenge. *Elem Sci Anth.* **2019,** *7*, (1).

8. Fox, T. A.; Hugenholtz, C. H.; Barchyn, T. E.; Gough, T. R.; Gao, M.; Staples, M., Can new mobile technologies enable fugitive methane reductions from the oil and gas industry? *Environmental Research Letters* **2021,** *16*, (6), 064077.

9. Hanley, J. A.; McNeil, B. J., A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* **1983,** *148*, (3), 839-843.

10. Obuchowski, N. A., ROC Analysis. *American Journal of Roentgenology* **2005,** *184*, (2), 364-372.

11. Young, M.; Fan, W.; Raeva, A.; Almirall, J., Application of Receiver Operating Characteristic (ROC) Curves for Explosives Detection Using Different Sampling and Detection Techniques. *Sensors* **2013,** *13*, (12), 16867-16881.

12. Receiver Operating Characteristic. https://en.wikipedia.org/wiki/Receiver_operating_characteristic.

13. https://en.wikipedia.org/wiki/Capping_inversion.

END OF REPORT