HISTORICAL CANADIAN FUGITIVE EMISSIONS MANAGEMENT PROGRAM ASSESSMENT

Prepared for: Petroleum Technology Alliance of Canada (PTAC)
January 30, 2017

GREENPATH ENERGY LTD

A C K N O W L E D G E M E N T S

The data provided by member companies of the Petroleum Technology Alliance of Canada (PTAC), as well as non-PTAC companies, was a key success factor in accomplishing the study. Guidance from Sean Hiebert of ConocoPhillips made this project possible during the condensed timeframe.

GreenPath Energy Ltd. would also like to thank PTAC for its funding contributions to make this project possible.

P R O J E C T   T E A M

The project team consisted of Michael D'Antoni, Joshua Anhalt and Samantha Clarke of GreenPath Energy. Bill Tubbs (formerly of Spectra Energy and ICF) was subcontracted to the project given the tight timelines of the project.

Any opinions, findings, and conclusions or recommendations expressed in this report are those of the author(s) and do not necessarily reflect the views of the reviewers or their agencies.

Project guidance was provided by:

Sean Hiebert, ConocoPhillips
Brad Johnson, Cenovus Energy
Nadia Monaghan and Kris Manipon, Encana
Greg Unrau, Repsol

Peer review of the document was provided by:

Jackson Hegland, Modern West Advisory
Don D’Souza, Ministry of Natural Gas Development

ABOUT GREENPATH ENERGY

GreenPath Energy Ltd. is North America's premier emissions management service provider for the oil and gas and petrochemical industries, specializing in emissions measurement and reduction solutions. This includes equipment inventory collection, leak detection and repair (LDAR) for fugitive emissions, and methane emissions reduction project development. Our technical expertise and diverse experience in emissions management ensures that we provide clients with solutions which allow for efficient use of capital while achieving significant emission reductions and regulatory compliance.

Our expertise in building 'best practice' fugitive and vented emission management solutions has been developed over the past nine years through our extensive instrumentation backgrounds and by using the best available technology. We engage regularly with government, regulatory bodies, industry associations and technology providers to ensure we are at the leading edge of solutions for emissions management programs.

EXECUTIVE SUMMARY

Throughout North America, federal and provincial or state level governments are developing methane emissions regulatory frameworks to be implemented in the coming years with a target of achieving a 40 to 45% reduction in methane emissions from the upstream oil and gas industry by 2025. Alberta has committed to a 45% reduction by 2025, with enhanced leak detection and repair likely to be one of the first policy frameworks which will be used to help achieve this objective.

The three most critical components to a Leak Detection and Repair (LDAR) program are facility inspection frequency, measurement standards, and repair requirements. To date, there is a significant gap in evaluating the current fugitive emissions management practices of the upstream oil and gas industry. The 2014 CAPP Update of Fugitive Equipment Leak Factors by Clearstone Engineering (2014 Clearstone Study) was designed to update leak factors for the National Inventory Report on Greenhouse Gases,[1] was based on a significantly smaller data set (approximately 120 facilities) and did not examine the effect of multiple inspections in different years on leaks at facilities. This document – the Historical Canadian Fugitive Emissions Management Program Assessment (FEMP Assessment) – is the first step to fill this gap.

The FEMP Assessment attempts to evaluate current industry LDAR practices, measurement methodologies, reporting mechanisms, leak and vent categorizations, and repair rates from a significant sample of Canadian oil and gas producers. It is the first study of its kind to closely evaluate the performance of these programs and methodically analyze the data to determine the primary sources and associated volumes of leaks and vents from upstream operations. This evaluation primarily serves to more accurately define fugitive emission management practice in Canada, and as the foundation for policy-makers to systematically evaluate regulatory options with regards to LDAR practices. The key output from this study is a large dataset on Canadian LDAR which can be expanded and further analysed at limited incremental cost and made available to stakeholders. GreenPath obtained fugitive emission management data from 14 different companies, accounting for more than 1,000 facilities, representing more than 1,200 different facility inspections, which detected over 10,000 leaks.

The data showed that, in general, existing FEMP programs have been targeting facilities with the highest leak rates. However, the sample size on small facilities (well-sites, batteries without compression, etc.) is too small to draw definitive conclusions on the leak rates by facility.

[1] http://www.capp.ca/publications-and-statistics/publications/238773

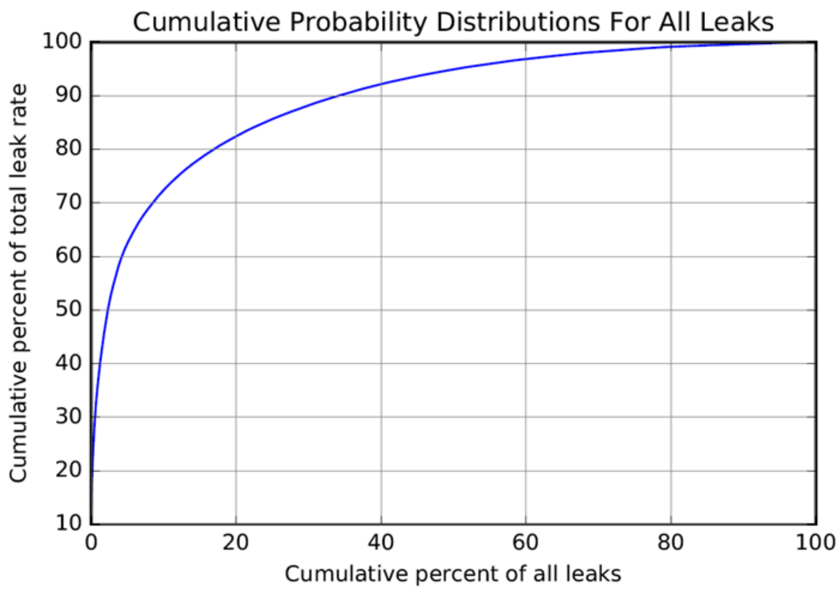
CONCLUSION #1: INCONSISTENT HISTORICAL LDAR DATA

The most important finding of the study is the need for industry to capture consistent, credible data: good data drives good decisions. As the industry moves forward to address the challenge of reducing methane emissions, the capture of concise, consistent, credible data will be critical to develop regulations which minimize cost to industry while maximizing methane reductions. To that end, GreenPath Energy has developed a standardized data capture template best practice recommendation which should be considered in go-forward FEMP programs. This template is attached in Appendix E. This data capture template has also been vetted by several LDAR subject experts, and has been deemed to contain the information required to produce credible, and consistent data sets, which will assist industry in drawing valuable conclusions from future data collected.

A key gap in the report is the lack of leak repair data. This study captures a wide array of data on leaks at upstream oil and gas facilities, but does not comment on repair effectiveness. The study also identifies the opportunity for improved consistency in measurement and reporting mechanisms for fugitive emissions. An investment in these opportunity areas will build a complete profile of fugitive emissions management practices that will improve industry performance and will ultimately support government's objective of achieving significant methane emissions reductions in the upstream oil and gas industry in Canada.

CONCLUSION #2: DATA ANALYSIS SHOWS VALUABLE TRENDS

Figure 1: 80/20 Rule - Canadian LDAR Data



As depicted in the figure above, it appears that the 80/20 rule applies to Canadian LDAR data. Approximately 80% of the total leak volume comes from approximately 20% the total leak count. This information supports the notion that a multi-layered cost-effective LDAR program methodology may be viable as it would allow for high level screening of large emission sources followed up by focused on-site assessments/repair. Technologies are rapidly developing to cost effectively identify major emission sources, and when paired with existing LDAR practice may serve to increase overall cost-effectiveness of LDAR programs.

Within the data set, we can also see that historical Canadian LDAR programs which focus on large complex facility types (Gas Plants, Gas Gathering Systems including Compressor Stations, and Oil Batteries) appear to be effectively targeting the highest probable leak rate facilities. This can be shown in Figure 2 below. Another important conclusion that can be drawn from the current data set is that smaller/less complex facility types (wellsites) are less leak prone. The one limitation on this conclusion is the small sample size related to wellsites and smaller facilities.

Figure 2: Average Leak Rate by Facility Type (CFM)

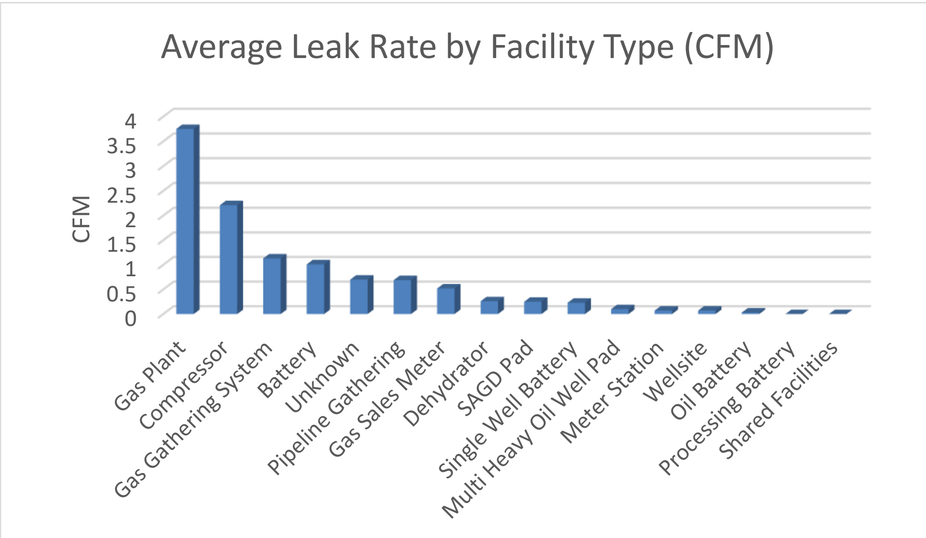


Table 1: Average Leak Rate by Component Type

Component Type	Number	Total Rate (CFM)	avg/leak
Connector	6539	573	0.09
Control Valve	2044	215	0.10
Valve	661	47	0.07
Unknown	431	178	0.41
Open Ended Line	249	172	0.69
Pressure Regulator	91	18	0.20
Pressure Relief Valve	52	192	**3.69**
Total	**10067**	**1395**	**0.14**

As shown in Table 1 above, the most common leaking components are connectors and control valves, but on average these leak point sources are small in volume. On the other hand, pressure relief valves are seldom found leaking, but when they do leak, could result in significant leak volumes. Appendix B provides more detail on the frequency distribution of leak volumes by component type.

NEXT STEPS

Tools and methods have been built to analyse large volumes of data quickly and consistently. As additional historical and future LDAR data becomes available, more detailed analysis can be undertaken.

TABLE OF CONTENTS

TABLES AND FIGURES

ACRONYMS AND KEY TERMS

AER	Alberta Energy Regulator
BMP	Best Management Practice
CAPP	Canadian Association of Petroleum Producers
CFM	Cubic Feet Per Minute
CO2e	Carbon Dioxide Equivalent
CSV	Comma-separated values
DLS	Dominion Land Survey of Canada
EPEA	Environmental Protection and Enhancement Act (Alberta)
FEMP	Fugitive Emission Management Plan
GHG	Greenhouse Gas
IPCC	Intergovernmental Panel on Climate Change
LDAR	Leak Detection and Repair
LSD	Legal Subdivision (smallest units in DLS)
NTS	National Topographic Service
OGI	Optical Gas Imaging
PTAC	Petroleum Technology Alliance of Canada
QA	Quality Assurance
QC	Quality Control
UOG	Upstream Oil and Gas
UWI	Unique Well Identifier
VRU	Vapour Recovery Unit

INTRODUCTION

GreenPath Energy Ltd. (GPE) of Calgary, Alberta was contracted by the Petroleum Technology Alliance of Canada (PTAC) to perform a desktop analysis of the existing Canadian Fugitive Emissions Management Plans, as well as leak detection and repair (LDAR) data.

Participant companies accounted for approximately 50% of all wells and facilities and approximately 33% of Alberta's barrel of oil equivalent (BOE) production. It should be noted that the data in the report does not represent all leak detection and repair data from these operators as some data could not be obtained or could not be integrated into the data set due to format (PDF instead of XLS or CSV files) and limited data processing budget. For example, in one producer's data set, only one sixth of the data could be processed, but that sixth processed represented 2,300 emission records, suggesting that the data set could be expanded considerably with additional resources/budget.

This study provides a robust data set upon which an analysis of LDAR practices in Canada can be evaluated. The objective was to provide qualitative data on current leak detection and repair practices in Canada and an analysis of existing data, which could be used to form an assessment of the effect of different inspection frequencies on emissions on leak rates from facilities. This data could be used to help inform policy-makers on the effectiveness of potential regulatory prescriptions regarding LDAR in Canada.

BACKGROUND

Leak detection and repair practices in the oil and gas industry have been evolving significantly over the last decade. The main driver behind the evolution of leak detection and repair practice has been technology such as optical gas imaging (OGI). Pioneered by FLIR Systems of Sweden (which introduced the Gas Find IR Camera in 2004), OGI can visualise methane and other gaseous hydrocarbons on oil and gas sites. Prior to the introduction of OGI, standard leak detection approaches required the use of toxic vapour analyzers (the “Method 21” approach), a time-consuming process.

The use of OGI as an alternative work practice (AWP) for leak detection was approved by the US EPA in 2008.[2] Canada was an early adopter of OGI for leak detection, as OGI was integrated into the Canadian Association of Petroleum Producers Best Management Practice for leak detection and repair in 2007 (CAPP BMP)[3]. There have been recent increases in regulatory requirements for leak detection activities in several states (Colorado, Wyoming, Pennsylvania, Utah, Ohio and California), as well as the US EPA New Source Pollution Standard (NSPS) and Bureau of Land Management (BLM) regulations.

Studies such as the Economic Analysis of Methane Emission Reduction Opportunities in the Canadian Natural Gas Industry[4] (ICF Study) and the Carbon Limits Study Quantifying Cost Effectiveness of Systemic Leak Detection and Repair Programs Using Infrared Cameras [5] (Carbon Limits Study) have framed leak detection and repair programs as highly cost-effective methods of reducing methane emissions from upstream oil and gas.

[2] Timeline of Optical Gas Imaging Regulation in United States and Europe. http://www.flir.com/uploadedFiles/Automation/Resources/Timeline-of-OGI-Regulations-January-2015.pdf

[3] Canadian Association of Petroleum Producers, Best Management Practice for Fugitive Emissions Management http://www.capp.ca/publications-and-statistics/publications/116116

[4] ICF, Economic Analysis of Methane Emission Reduction Opportunities in the Canadian Oil and Natural Gas Industries, https://www.pembina.org/reports/edf-icf-methane-opportunities.pdf

[5] Carbon Limits Quantifying Cost-Effectiveness of Systemic Leak Detection and Repair Programs Using Infra Red Cameras, http://www.catf.us/resources/publications/files/Carbon_Limits_LDAR.pdf

DATA

Leak detection and repair (LDAR) programs using optical gas imaging (OGI) have been in Canada since the mid-2000s by various producers. In 2007, the Canadian Association of Petroleum Producers (CAPP) *Best Management Practice for Fugitive Emissions* was published to provided best practice for leak detection and repair practice across Canada, and was subsequently integrated into the regulatory framework in Alberta and British Columbia in the flaring and venting guidelines of the AER (2011-35), and *OGC Flaring and Venting Guidelines* (2013). Data has been collected from producers with internal LDAR programs and third party LDAR service providers. The CAPP BMP allows for significant flexibility, including procedures for the collections and management of data. However, this data has not been previously compiled and analyzed in a consistent manner. This study uses advanced analytic techniques to compile dissimilar data sources to develop a more consistent dataset on emission leak detection and repair activities from 2007 to 2016.

In relation to the 2014 Clearstone Study, the data set analyzed in this study represents a larger number of facilities and attempts to determine if patterns emerge among facilities that have been inspected multiple times. In relation to the Carbon Limits study, a wider array of data has been analyzed from multiple third party LDAR providers as well as producers with LDAR programs. The Carbon Limits study data set includes data from American operations, whereas this study includes only Canadian data.

This data set is also focused on leaks as opposed to vents; the initial data request that went out to industry partners requested leak data, but some companies provided LDAR data with both leaks and vents. The data set within this study should not be used to analyze venting, as LDAR surveys are designed to focus on the detection and quantification of leaks.

STUDY OBJECTIVES

There were three main objectives within the study:

- Analyse all available historical Canadian FEMP data.
- Determine the occurrence and importance of "super-emitters" in relation to fugitive emissions, and determine what percentage of leak volume comes from a specified percentage of leaks (e.g. does the 80-20 rule apply?).
- Provide a significant, yet anonymous dataset which could be analyzed further to draw meaningful and influential conclusions (such as LDAR program effectiveness and changes in inspection frequency on leak rates)

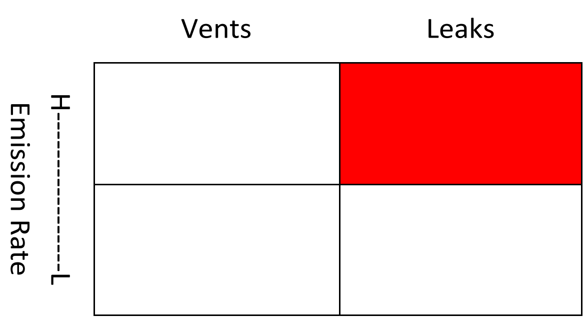
S U P E R - E M I T T E R S

The concept of "super-emitters" is a new term which has recently entered the lexicon in discussions of emissions from oil and gas operations from academic papers published in the United States by the Environmental Defense Fund (EDF), and into public discourse on methane emissions largely due to FLIR Systems' videos of the methane leak at the SoCal Aliso Canyon natural gas storage facility near Los Angeles, California.

The term was further expanded upon in academic papers.[6] For the purposes of this paper, we have defined a super-emitter as: "an individual component emitting greater than 6 SCFM (standard cubic feet per minute) or a facility with emissions greater than 30 SCFM". The existence of "super-emitters" helps resolve a disconnect in academic literature on methane emissions in the oil and gas sector, where top-down surveys (aerial, drive-by mass spectrometer) do not match with bottom-up emission factor estimates.

The concept of super-emitters is not limited to unintentional releases of methane, and thus includes sources which are designed to release methane such as tank venting or venting from compressor wet seals or other high emission sources on site.

Figure 3: Graphical Description of Super-Emitter Concept



In relation to leak detection and repair programs, the super-emitters of interest are in the top right corner with high leaking emission rates. High-volume vented emissions, shown in the top left, have other technology and best practice solutions for mitigation.

[6] Methane Leaks from Natural Gas Systems Follow Extreme Distributions
http://pubs.acs.org/doi/abs/10.1021/acs.est.6b04303
Reconciling divergent estimates of oil and gas methane emissions.
http://www.pnas.org/content/112/51/15597.abstract
Aerial Surveys of Elevated Hydrocarbon Emissions from Oil and Gas Production Sites,
http://pubs.acs.org/doi/abs/10.1021/acs.est.6b00705,
Methane Emissions from the Natural Gas Transmission and Storage System in the United States
http://pubs.acs.org/doi/pdf/10.1021/acs.est.5b01669

METHODOLOGY

This section outlines GreenPath Energy's approach and methodology used to complete this study including data collection, data processing, data analysis, and production of results.

DATA COLLECTION

Data was provided by 14 oil and gas producers in Canada, with most of the data pertaining to leak detection and repair within Alberta. Individual meetings held with several producers indicated their interest in providing data; the response was positive overall, though not all data could be processed prior to the publication of this study due to the large volume and inconsistent format of data contributed.

The Fugitive Emissions Management Plans (FEMP) outline company policies about inspection frequency and repair practice when and where leaks are detected. Appendix A: Fugitive Emission Management Plans provides a tabular comparison of FEMPs made available to GreenPath.

Data was provided in several formats; this in addition to variations in data capture process meant that not all data provided could be used in a meaningful fashion. Many files received were in PDF format, unable to be fully "scraped" for their contents to be integrated into the data set. In the case of other firms, third-party service providers did not provide quantified or estimated emissions rates, making the data of limited value in determining leak rates. Over 200 hours were spent gathering and processing data, and developing tools to pull data from disparate sources into a common data set that would allow for comparisons. The tools developed for this project are flexible, such that GreenPath can add large volumes of additional data to the data set at minimal cost.

Firms engaged in the study have signed non-disclosure agreements with GreenPath Energy. As such, all data within the report is confidential with respect to company name, facility location, and fugitive emission management plans.

In relation to the FEMP operating companies, a letter code has been used to replace company names. Only GreenPath Energy is privy to the code which attributes FEMPs to a given company.

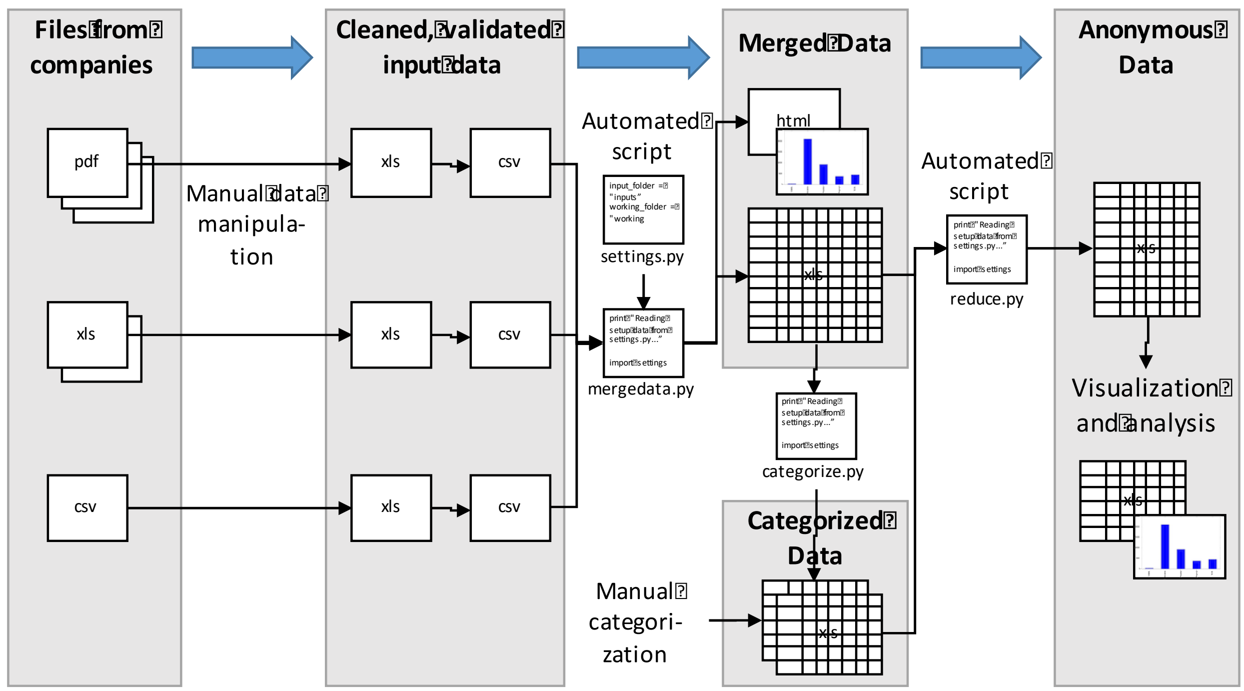
DATA PROCESSING

GreenPath Energy developed a customized set of data analysis and processing tools for this study, as necessary to implement a sufficiently robust process to combine the non-uniform data received from multiple sources.

The data processing involved the following tasks and is described in the figure below:

1. Checking and validating input data (mostly manual).
2. Identification and labelling of data (automated with manual supervision).
3. Combining data into one data file (automated).
4. Categorization of data (manual with assistance from a text classification algorithm).
5. Removal of company-sensitive and unnecessary information (automated).

Figure 4: Data Processing Diagram



Participating companies submitted leak detection and repair data in three different document formats: Adobe PDF (pdf), Microsoft Excel (xls/xlsx), and Text file (csv).

In some cases, data from multiple leak detection surveys was already combined into one fileand in others we received multiple files from a company. The table below summarizes the number of files received from each company, the number loaded, and the total number of records extracted from each company.

Table 2. Inventory of Source Data Files

Company	Number of files provided	Number of files loaded	Number of emission records processed
P	2	2	919
N	231	37	2,380
K	11	11	3,667
T	4	4	31
F	83	46	754
Q	1	1	754
L	16	16	2,109
H	2	2	1,020
A	25	25	384
S	3	3	768
M	1	1	3,110
E	1	1	525
TOTAL	**380**	**149**	**16,421**

A large portion of the overall data provided by operators was loaded and processed. The main exceptions were many files from Company N and Company F, provided in PDF. Although we identified tools to extract data from tables in a PDF document, it is cost-prohibitive to extract the entirety of the data provided in this way, so a limited amount of data was used from the sample.

Early on, we decided to automate the data process using the Python scripting language.[7] One of the critical advantages of automating data processing is that it can be re-executed numerous times, allowing new data (provided sporadically over the course of the project) to be added and easily reprocessed with the entire data set. At each step of the process, we reviewed and checked the outputs of the script to ensure it was working correctly.

MERGING DATA

One of the biggest challenges was to identify similar data in each source file and combining it into one logical data file. Leak surveys carried out at different times by different service providers, with individuals recording unique data types and labels are put into a uniform format so that analysis can be undertaken.

For example, site locations were described in nine different ways. This resulted in the development of a reliable system of combining all columns that contained similar data by establishing 'dictionaries' for each data type common to most surveys. Leak rates were recorded in seven different units with different combinations of units reported by operators and leak detection and repair firms.

Although each leak survey report contained different data, we chose a set of data which appeared to be most useful for the analysis and then attempted to match data in each report. We manually added key fields such as 'Company', 'Year', 'Province' and 'Source Filename', which were not included in the survey reports to maintain company confidentiality.

[7] Python (http://python.org) is a general-purpose, high-level programming language that is excellent for customized data processing and analysis tasks.

Table 3: Desired Data Types and Number of Records Found

Data Type	% of total	Number of records filled
Company	100	16,508
Year	100	16,508
Location	100	16,508
Survey Date	99	16,505
Source Filename	99	16,421
Intentional	99	16,465
Province	92	15,220
Building/Process Unit	88	14,532
Component Category	83	13,807
Leak Rate [CFM]	82	13,649
Emission Description	72	12,000
Facility Name	70	11,713
Leak Rate [e3m3/year]	63	10,418
Tag ID	62	10,377
Component	55	9,218
Area	54	9,045
Quantification Method/Estimated	39	6,515
Gas Type	35	5,910
Leak Notes	33	5,603
Facility Type	19	3,264
Leak Rate [L/min]	12	2,129
Leak Rate [MmBTU/ year]	12	2,029
Leak Rate [e3m3/day]	12	2,017
Repair Date	12	2,036
Successfully Repaired	8	1,482
Instrument Type	3	522
Leak Rate [SCFY]	2	378
Vent Rate [e3m3/day]	0	7

GreenPath manually reviewed each input data source, identifying matches between the data types provided and the desired data types, then automating the merging process by building a Python dictionary to act as a correspondence table between input data labels and desired data types. This allowed for an automatic merging of the data from dissimilar sources.

The following code extract shows an entry in the column label dictionary to illustrate how the script dealt with inconsistencies in source data labelling. In this example, it identified nine possible data as the site location and assigns the desired type labelled 'Location'.

```
'Location': ['Location', 'Site', 'Site LSD', 'siteName', 'Site Name (LSD)',
'Site Name', 'Facility Land Location', 'Facility LSD (xx-xx-xxx-xx-WxM)', 'LSD'],
```

IDENTIFYING UNIQUE SITES AND LOCATIONS

To track changes in the leaks detected over time (LDAR effectiveness), we needed to link surveys carried out at the same site at different times, which proved difficult because of inconsistencies in reporting facility names and facility land locations. To solve this problem, we wrote a script that identifies a potential land location code (either LSD or NTS) and converts it to a standardised version of the code that can be used to identify a unique site.

The table shows examples of land locations from survey reports, and the standardised versions after conversion by the algorithm.

Table 4: Examples of Standardized Site Location Codes

Standardized Codes	Examples of Data Reported
'02-08-017-13W4'	'02-08-017-13-W4 ', '2-8-17-13w4'
'02-17-063-08W6'	'02_17_063_08W6M', '2-17-63-8-W6'
'09-27-079-17W6'	'9-27-79-17-W6', '09-27-079-17'
'10-09-073-13W6'	'10-9-73-13-W6', '10_09_073_13 W6M'
'10-10-071-08W6'	'10-10-071-8W6', '10-10-071-8W6M', '10_10_071_08 W6M'
'15-08-056-23W5'	'100/15-8-56-23 W5'
'A-029-H/093-P-09'	'A-029-H 093-P-09', 'A_029_H_093_P_09'
'B-100-B/093-P-08'	'B-100-B 093-P-08', 'B_100_B/093_P_08'
'C-067-K/094-O-08'	'C- 067-K/094-O-08'
'D-073-B/093-P-08'	'D-073-B 093-P-08', 'D_073_B/093_P_08'

By applying the standardisation algorithm, we reduced the number of apparently unique sites by 192 by identifying matches between site names where the name was reported inconsistently. It was noted that in some cases, these represented facilities where ownership had changed hands, facilities with multiple owners during the study period represented only one site.

Further processing and analysis of Company N's data, who acquired several facilities from other operators may increase the number of sites with multiple owners over time.

UNITS OF MEASUREMENT

Leak rate data was reported by service providers in a variety of units including:

- British thermal units per year (MMBTU/year)
- Litres per minute (L/min)
- Cubic feet per minute (CFM)
- Standard cubic feet per year(SCFY)
- 1000 cubic metres per day (e3m3/day)
- 1000 cubic metres per year (e3m3/year)
- Cubic meters per hour (m3/hour)

DATA ERRORS

Although we were not able to fully validate all the data we received, we are aware of inconsistencies, with the most obvious example from a survey report carried out in 2011.

The table below shows a subset of the data in this report.

Depending on which figures you choose to use to convert to standard cubic feet per minute (CFM), it appears to show multiple leaks in excess of 50 CFM at the same site. Likely the true reading was the l/minute figure but tracing back to the leak survey document, it was unclear what the direct reading was, the only way to confirm volume would have been to obtain the original leak videos and make a qualitative assessment of volume.

Considering the inconsistencies between figures reported for litres per minute (L/min) and CFM, this seems to be a reporting error. We removed this survey from the data. No other data was removed from the dataset. An evaluation matrix was created to flag data where leak rates for the same leak were incongruous. Only the data from this survey was flagged for exclusion.

Table 5: Leak Data with Inconsistencies

Leak Rate [MmBTU/ year]	Leak Rate [L/min]	Leak Rate [CFM]	Leak Rate [SCFY]	Leak Rate [e3m3/ year]	Leak Rate [e3m3/ day]	Leak Rate [m3/ hour]	Converted Leak Rate [CFM]	Year Surveyed
64,493	0			1,826	5.00		123	2011
27,087	10			767	2.10		52	2011
32,247	15			913	2.50		61	2011
128,987	10			3,653	10.00		245	2011
32,247	0			913	2.50		61	2011
128,987	0			3,653	10.00		245	2011
128,987	10			3,653	10.00		245	2011
		200		2,977			200	2014
		100		1,488			100	2014
		150		0			150	2016

SITE TYPE CLASSIFICATION

As part of the analysis, we wanted to understand how leak frequencies and leak rates are determined by the type of site or facility.

Table 6: Number of Unique Sites by Province and Type

Site Type	AB	BC	SK	Unknown	Totals
Unknown[8]	528	30	44	6	608
Compressor	178	85	0	0	263
Gas Plant	45	15	0	0	60
Battery	56	9	0	0	65
Oil Battery	2	0	0	0	2
Wellsite (unknown commodity)	11	28	0	0	39
Meter Station	1	36	0	0	37
Compression	29	4	0	0	33
Gas Well	1	0	0	0	1
Gas Gathering System	12	0	0	0	12
SAGD Pad	11	0	0	0	11
Multi Well Heavy Oil Pad	5	0	0	0	5
Processing Battery	0	3	0	0	3
Single Well Battery	3	0	0	0	3
Single-Well Gas Battery	2	1	0	0	3
Pipeline Gathering	0	2	0	0	2
Dehydrator	2	0	0	0	2
Gas Multiwell Group Battery	0	0	0	0	0
Gas Sales Meter	0	1	0	0	1
Shared Facilities	0	1	0	0	1
Totals	**886**	**215**	**44**	**6**	**1151**

There were two major challenges in allocating leaks to different facility types. First, a large volume of the data did not identify based on DLS or NTS co-ordinates; for example, “Large Rock Facility” cannot be placed into DLS or NTS co-ordinates and/or matched in the AER or BCOGC data sets. The naming system as well as potential errors in the records of DLS and NTS locations limited the ability of GPE to classify all facilities. In cases with provided DLS or NTS co-ordinates, many co-ordinates matched with multiple permitted types.

Some facilities matched between 4 and 9 different permitted types with the AER data set. Where multiple matches were found, the “largest” facility type was assigned. For example, Gas Plant > Gas Gathering System> Compressor Station > Battery > Well.

[8] Note almost all of unknown facilities are likely assets with compression on site (compressor or gas plant).

If a facility match produced all four facility types, the facility would be manually assigned the category "Gas Plant". The challenges with facility identification resulted in 31% of the leak volume not being attributable to a specific facility type.

Additionally, the facility categories employed by the AER, are too diverse to draw meaningful categories with 37 potential types of facilities. In further studies GreenPath has suggested that facilities be classified in the following matrix: [well, multi-well, battery, compressor, plant], commodity [oil or gas primary] and service [sweet or sour].

EMISSION TYPE CLASSIFICATION

At a high level, emissions detected as part of a FEMP program can be put into two primary categories, leaks (unintentional releases of methane) and vents (intentional releases of methane). Table 6 below shows the typical classification of leaks versus vents in most LDAR data. Within the data, compressor seal venting (reciprocating compressor packing venting and centrifugal compressor seals) was inconsistently categorized as a leak or a vent. GreenPath has manually categorized emissions from compressor seals as vents.

Table 7: Examples of Leaks versus Vents in LDAR Data

Leak (unintentional)	Vent (intentional)
Loose connection	Pneumatic Controller Vent
Valve passing	Pneumatic Pump Vent
Valve diaphragm	Dehydrator Vent
Pressure Release Valve Passing	Compressor Rod Packing Vent
Valve stuck open	Wet-Seal Venting
Unintentional hole in pressure system	Casing head gas venting
Venting from tank (if VRU down)	Venting from Tank (if no VRU in place)
Non-routine venting	Unlit flare

Within the data, concepts of leaks or vents are not universally categorized as such. For example, one LDAR firm within their own data set used the following terms to differentiate leaks versus vents: with 'Intentional' as the column heading, valid responses were: *True, False, Yes, No, I, U, I, Leak, Vent*. Within the data set were some inconsistencies in terms of how leaks or vents were defined.

Some firms did not inventory or quantify vents and thus the difference in definition of leaks and vents could result in some material discrepancies between two firms examining the same site. GreenPath did not correct these potential errors except in relation to compressor seal venting. In terms of the data, the use of uncontrolled documents and forms, small changes such as Vent vs vent, or Leak vs leak would result in different categories. GreenPath developed a matrix to translate different leak and vent definitions into a common framework.

Table 8: Leak and Vent Categorization

Original Classifier	GreenPath Classification
I	Vent
UI	Leak
I	Vent
I	Vent
Leak	Leak
Mandatory Leak	Vent
Mandatory No Emission	Vent
Mandatory Vent	Vent
No	Leak
No Leak	NIL
No Leaks	NIL
No Leaks Found	NIL
U	Leak
UI	Leak
Vent	Vent
Vent (intentional emission)	Vent
Yes	Vent
FALSE	Leak
TRUE	Vent

Separating out a simple binary data identifier proved to be a complex task that required significant manual manipulation of the data. Certain data sets had "deep-dives" undertaken to examine irregularities.

A manual examination of the top ten leaks and vents by volume showed that three were misclassified (see Appendix C: QA/QC of the Largest leaks and Vents). A manual examination of the top ten leaks showed that four "leaks" were, in fact, tank vents. The inconsistencies of the binary choice highlight the issue of a lack of standardization of leak and repair data.

A further random sample of 100 leak records were selected to determine if the classification as leak or vent were accurate. From this sample, 95 percent were accurately categorized leaks or vents. Errors in categorization were generally related to compression seals and tank vents.

All components categorized as "compressor seals" have been re-categorized as vents. Tank vents do not correspond to a specific component type and thus could not be rapidly grouped as leaking or venting, as the description of "tank vent" was often contained within a freeform text field. In future analysis, a category for "tank venting" should be created to prevent this type of error.

Leak detection and repair surveys are driven by exceptions, as the fundamental nature of these surveys is to report back on unintentional releases of methane. One potential gap in the data set is the number of leak-free sites. The various data systems manage leak-free inspections differently, with some internal programs generating a record of a leak-free site, where others retain the formerly leaking value as a "saved rate" – to demonstrate the cost-effectiveness of the emissions reduction program. Thus, there may be sites that were inspected, found to be leak free, but no evidence of the inspection is contained in the aggregated data set. Within the data set, there are 94 facilities with an annual emission record of zero, but an indication that the facility was inspected. A facility with zero emissions (no venting) could conceivably been inspected but not integrated into the data set.

COMPONENT TYPE CLASSIFICATION

Because of the large number of leak records, we combined processes: manually labelling over 800 records, then training a machine learning algorithm to recognize component types from the text data reported by the leak survey technicians.

We used a technique common to text data analysis called Multinomial Naïve Bayes classification. The algorithm was trained on roughly 600 labelled data points, then tested on the remaining 200 labelled points. Accuracy was found to be 88%, which was deemed satisfactory for this analysis.

The component-type categories that were generated were derived from the 2014 *Canadian Association of Petroleum Producers Report – Update of Fugitive Equipment Leak Factors* by Clearstone Engineering[9].

Table 9: Component Types

Component Type
Connector
Open-Ended Line
Pressure Regulator
Pressure Relief Valve
Control Valve
Valve
Compressor Seal
Unknown

A category of "unknown" was added where the component could not be accurately determined, flagging instances where the algorithm was not confident that one of the seven types of components could be determined. Human review of the data also showed several leak records where the type of leak could not be adequately determined. Further research could be used to classify "unknown" categories or generate a more exhaustive list of component categories.

Once trained, the algorithm could be used to predict the classification of the component type for all remaining leak records. As a precaution, we manually checked over 233 records where the predictions of the machine learning algorithm were uncertain, and a further 49 randomly selected records (Table 10: Random Sample Comparison of Human versus Machine Learning Component Categorization).

The randomly-selected categorizations showed positive results, with more than 75% of the components accurately categorized. Among common components such as connectors or open-ended lines, the machine learning algorithm correctly predicted over 80% of the correct component type.

The algorithm was less accurate with less common component types, and further manual categorization was undertaken on compressor seals. Control valves and pressure regulators are relatively rare within the data set (see Appendix B). The machine learning algorithm was quite effective in properly classifying these component types when a closer examination of these types was undertaken.

Table 10: Random Sample Comparison of Human versus Machine Learning Component Categorization

	Human Classification	Machine Learning Prediction	Total Samples	% Correct
Compressor Seal	3	1	3	33%
Connector	27	23	27	85%
Control Valve	4	4	4	100%
Open Ended Line	12	10	12	83%
Pressure Regulator	1		1	0%
Unknown	1		1	0%
Valve	1	0	1	0%
Total	**49**	**38**	**49**	**78%**

[9] http://www.capp.ca/publications-and-statistics/publications/238773

EQUIPMENT TYPE CLASSIFICATION

Equipment type was a reasonably common classifier within most data files – for example: compression, dehydration, filter/separation. We provided the capability to label equipment types in the data but did not feel it would have added sufficient value; project budgets and timelines did not allow us to complete the classification of equipment types.

LEAK VOLUME CLASSIFICATION

Limited standardization exists in relation to reports generated from leak detection and repair surveys between companies. Leaks are measured in several different formats ranging from MMBTU/minute to L/minute, standard cubic feet per minute (SCFM or CFM), as well as e^3m^3/year and e^3m^3/day. CFM was chosen as the metric for leak volume, as the primary quantification device (the Bacharach Hi-Flow Sampler) reports in CFM.

Attempts were not made to standardize values to standard temperature and pressure as previous analysis has shown that these adjustments do not generate a material change in reported emissions; in addition, tracking down often unknown weather and barometric pressure for sites would prove problematic.

Inconsistencies in the reported leak volumes required additional classification for analysis purposes. One issue that was encountered was that leak data was periodically inconsistent. For example, one leak detection and repair company quantified leaks at L/minute, e3m3/year, and occasionally e3m3/day within the same report.

In some surveys, it appeared there were conversion errors in the reported leak volume data. For example, if the e3m3/year value was used to determine CFM, a leak rate of 245 CFM was generated; but if the litre per minute value was used, then 0.35 CFM was generated. It is most likely that the litre per minute figure was the actual measurement, as the high flow sample can also output in litres per minute in addition to CFM. These data points were excluded from analysis due to their potential to skew results.

To standardize leak volumes, a formula was developed which looks first to the CFM value, then the litres per minute value, then an e3m3/day value, and finally an e3m3/year value to translate units in CFM.

In addition to issues in converting leak rates to standard units, leak repair data did not uniformly state whether the result was estimated via the operator visualizing the hydrocarbon plume via OGI, or quantified using a Hi-Flow Sampler. An examination of individual data points can often suggest whether data was estimated or quantified. For example, the Hi-Flow Sampler has a maximum rate of 10 CFM, but becomes significantly less accurate above 5 CFM and can detect flow rates as low as 0.01CFM. Thus, rates above 10 CFM are generally estimated (exceptions are data that can be traced back to a calibrated volume bag or other method).

COMMERCIALLY SENSITIVE INFORMATION

To protect companies who provided data and to allow the results to be openly discussed by project stakeholders, we chose a policy to remove all company-specific and identifying information from the final merged data file and produce a reduced data file with only leak rate, survey date, and assigned categories.

To achieve this without aggregating the data, we replaced all site locations with a randomly-generated 'site key' and every leak record with a randomly-generated 'leak key'. The table below shows a section of this reduced data table providing the data types and a few data points.

Table 11: Reduced Data Table

Leak Key	Assigned Component Category	Assigned Emission Type	Assigned Equipment Type	Converted Leak Rate [CFM]	Province	Site Key	Site Type	Survey Date
Nm6nndnV	Unknown	Vent	Unknown	0	AB	05S5w2	Unknown	23/05/2013
5bz52vIB	Unknown	Leak	Unknown	0.35	AB	0IjjZp	Unknown	7/10/2011
QIBHqGX6	Open Ended Line	Vent	Unknown	10	AB	0IjjZp	Unknown	7/10/2011
9JqII63I	Open Ended Line	Vent	Unknown	5	AB	0IjjZp	Unknown	7/10/2011
y5mdD1zN	Open Ended Line	Vent	Unknown	0.52	AB	0IjjZp	Unknown	7/10/2011
ryCyjtJq	Connector	Leak	Unknown	0.12	AB	0IjjZp	Unknown	7/10/2011
LSs9ay4q	Connector	Leak	Unknown	0.08	AB	0IjjZp	Unknown	7/10/2011
…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…

There is currently a total of 14,764 data points in the reduced output data file. With the approval of PTAC and project participants, this anonymous data file has been shared with all stakeholders.

DATA ANALYSIS

Data has been analysed using a combination of Python notebook outputs and Microsoft Excel to parse the data into meaningful categories.

QA/QC

Throughout the process, data points have been checked for integrity. This review resulted in the discovery of a mismatch in leak rates within a survey; it should be noted that the error only related to one specific survey. This anomaly was only detected due to the very large volume results presented.

Samples within the data set have also been evaluated for consistency, the top ten leaks and vents have been evaluated and errors within rectified.

Examining the top 10 emitters also showed inconsistency in how LDAR companies and oil and gas producers determine whether compressor seals were leaks or vents; GreenPath has manually re-categorized those components correctly (see Appendix C: QA/QC of the Largest leaks and Vents).

A random sample of ten leaks were traced back through the merging process to original documents to determine if the data had been corrupted (leak volumes, categorization, etc.) and no errors were found. Further QA/QC via manual verification of data is recommended on the data set prior to distribution.

CONFIDENTIALITY

To maintain the confidentiality of participating companies, the data went through a rigorous process to remove identifying information. In the global data set, company names were replaced with a randomly generated key code. Only GreenPath Energy has access to the lookup table that links the key codes to the company and site location information and facility names.

FUGITIVE EMISSION MANAGEMENT PLAN REVIEW

In addition to the raw data from fugitive emission management plans, nine individual company Fugitive Emission Management Plans were provided to GreenPath Energy for analysis. All existing FEMPs are based upon the existing CAPP Fugitive Emission Best Management Practice.

The CAPP BMP framework allows for flexibility in directed inspection and maintenance approach. Some plans are based on a US EPA "Method 21" toxic vapour analyzer approach, whereas current practice commonly employs optical gas imaging. There is significant variation in inspection frequency, data collection, roles and responsibilities.

A tabular summary of the FEMP review is provided in Appendix A: Fugitive Emission Management Plans . The existing CAPP framework allows for flexibility in addressing fugitive emissions, but the variations in implementation and execution generated some of the challenges related to compiling the data set used in this study; as discussed in a following section. Not all FEMPs submitted were accompanied by useful leak data. For example, some companies did not report volumetric leak information.

RESULTS

The data has been compiled and analysed using the following charts and graphs to present leaks from upstream oil and gas emissions in Canada.

Figure 5: Total number of survey reports processed by year in this study

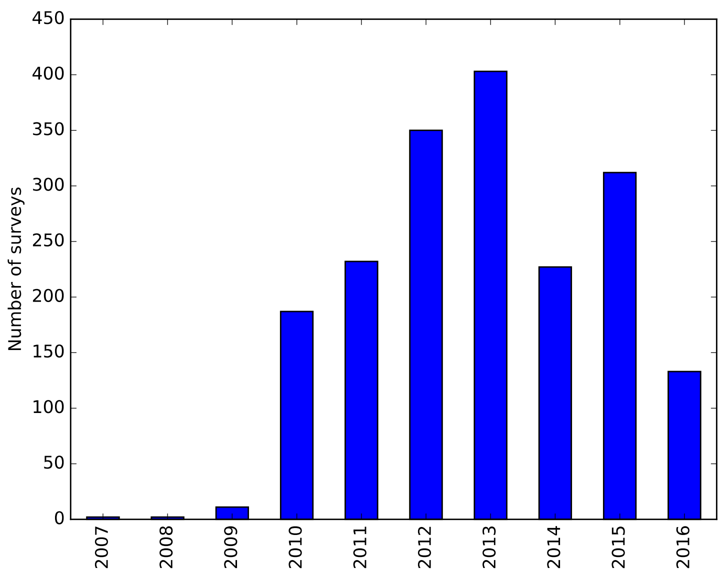
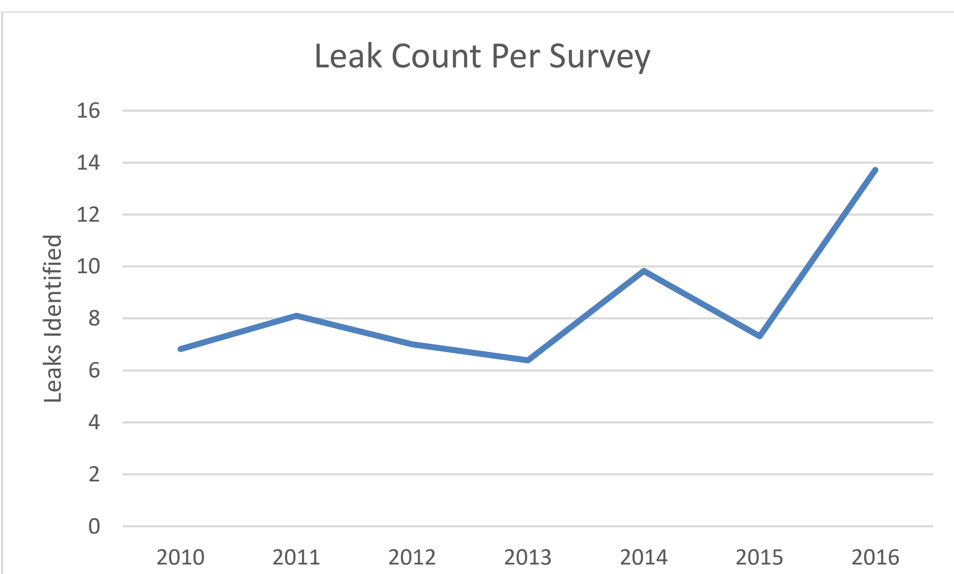
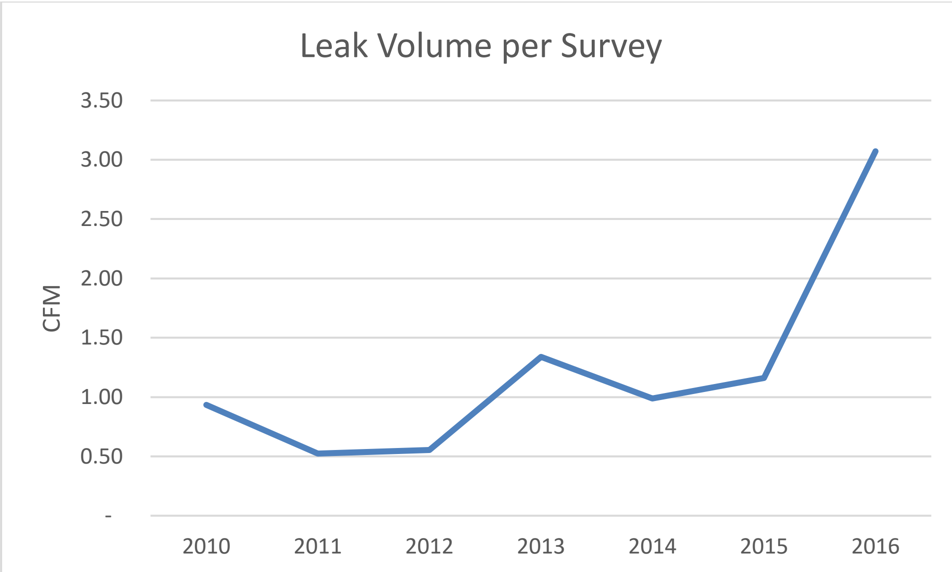


Figure 6: Average number of leaks detected by survey per year



There are several possible explanations for the increasing trend in leaks reported per survey. Increased use of the FLIR GF320 has increased the ease of spotting leaks with OGI; with decreased budgets, some firms have focused their LDAR efforts on known high-emission facilities.

Figure 7: Average leak volume per survey per year



Note: The spike in leak volume in 2016 is primarily attributed to the discovery of a substantial leak in 2016. The leak was found by GreenPath Energy, and estimated using a calibrated volume bag. This leak filled the calibrated volume bag in under two seconds; it is very important to note that the leak was resolved within two weeks.

Figure 8: Leak volume per survey (substantial leak removed)

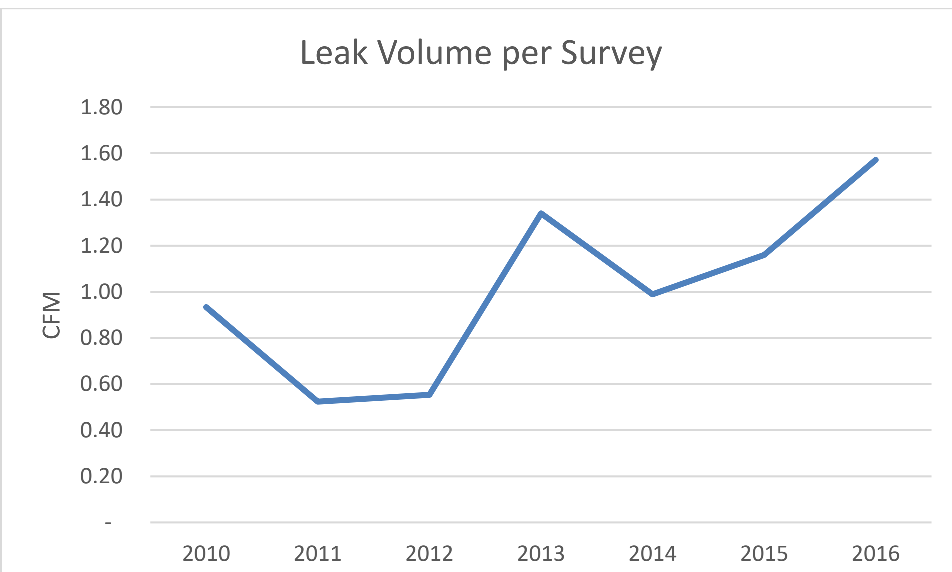
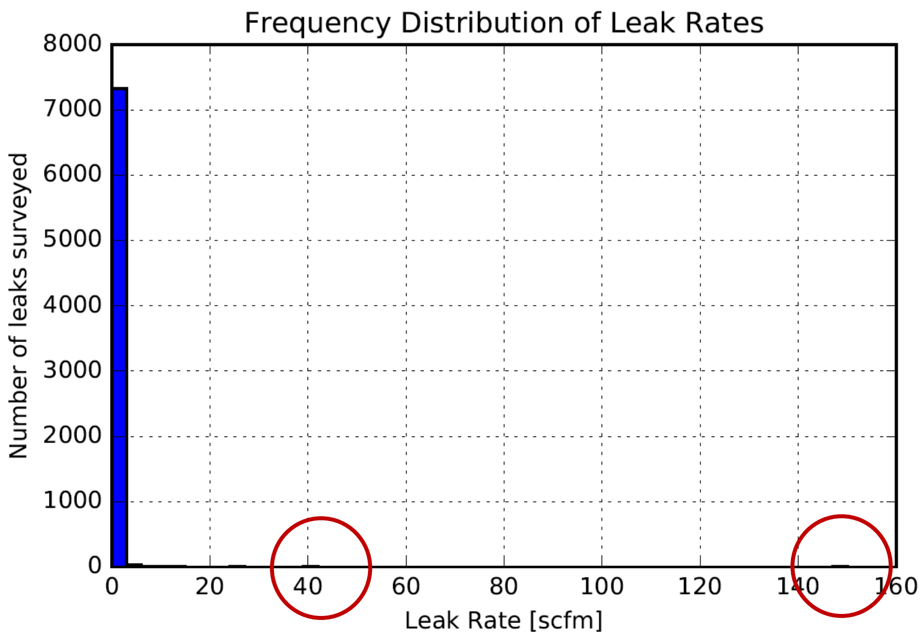
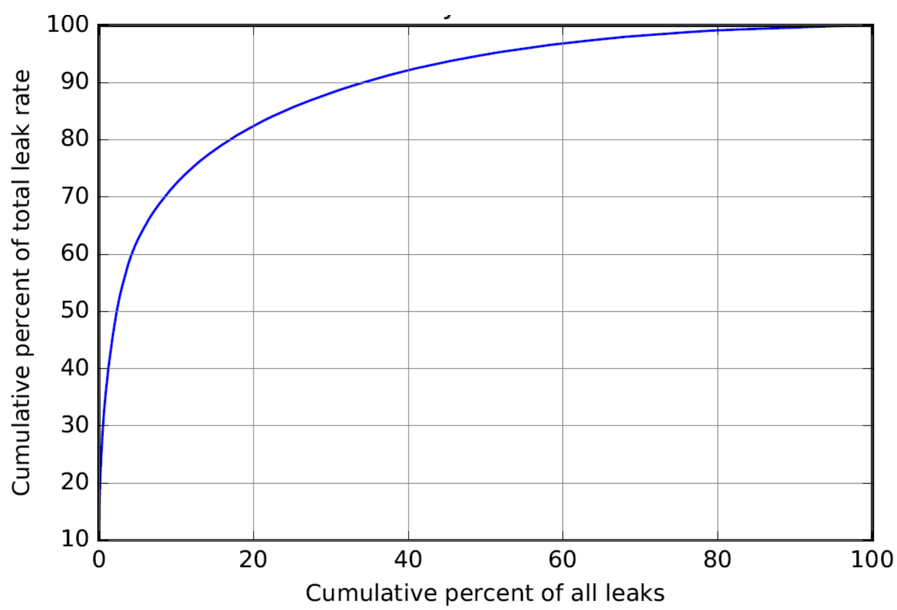


Figure 9: Frequency distribution of leak and vent rates



The frequency distributions of leaks show that the clear majority of leaks, but a small number of outliers raise the average emissions.

Figure 10: Cumulative probability distribution of all leaks



The cumulative probability distribution of all leaks shows that less than 20% of all leaks are responsible for 80% of all leak volumes. Thus, targeting components with high potential leak volumes is of significant benefit. A breakdown of the charts above by component types is available in Appendix B.

Table 12 illustrates that pressure relief valves are disproportionately responsible for large emissions. A leaking pressure relief valve is more likely to result in a super-emitter based on the data.

Human categorization is required to determine the “unknown” emitting components. Even with expert opinion some “unknown” components could not be categorized with confidence. The component type “open-ended line” was excluded as most of these were attributed to vents; improper categorization could skew results.

Table 12: Components with leak rates >6CFM

	Connectors	Pressure Relief Valves	Control Valves and Valves	Unknown
# of leaks >6CFM	13	3	18	3
% leaks >6CFM	0.2%	5.45%	0.53%	5.45%

Table 13: Leak count and rate by component type

Component Type	Number	Total Rate (CFM)	avg/leak
Connector	6539	573	0.09
Control Valve	2044	215	0.10
Valve	661	47	0.07
Unknown	431	178	0.41
Open Ended Line	249	172	0.69
Pressure Regulator	91	18	0.20
Pressure Relief Valve	52	192	3.69
Total	**10067**	**1395**	**0.14**

Table 14: Total leak rates by facility types (CFM)

Facility Type	Total Leak Rate	Grand Total
Battery	65.44	165.36
Compressor	653.26	1,289.43
Gas Gathering System	13.55	34.77
Gas Plant	224.82	326.82
Gas Sales Meter	0.52	1.14
Meter Station	2.67	8.19
Multi Heavy Oil Well Pad	0.51	0.51
Oil Battery	0.06	0.06
Pipeline Gathering	1.37	4.10
Processing Battery		7.73
SAGD Pad	2.74	2.74
Shared Facilities		0.91
Single Well Battery	0.68	0.68
Unknown	426.29	1,308.66
Wellsite	2.62	11.79
Grand Total	**1,395.08**	**3,165.16**

Table 15: Average leak and vent rates per facility[10]

Facility Type	Avg Leak Rate Per Facility (CFM)	Count of Facilities
Battery	1.01	65
Compressor	2.21	296
Dehydrator	0.26	2
Gas Gathering System	1.13	12
Gas Plant	3.75	60
Gas Sales Meter	0.52	1
Meter Station	0.07	37
Multi Heavy Oil Well Pad	0.10	5
Oil Battery	0.03	2
Pipeline Gathering	0.69	2
Processing Battery	-	3
SAGD Pad	0.25	11
Shared Facilities	-	1
Single Well Battery	0.23	6
Unknown	0.70	608
Wellsite	0.07	40
Grand Total	**1.21**	**1151**

[10] Note: Approximately 36% of the data could not be adequately classified by facility type.

DISCUSSION OF RESULTS

One of the major limitations of this study is the lack of consistency in relation to data collection and methodologies. For example, some firms inventory and measure vents in addition to leaks. Some firms quantify *only* leaks; others do not quantify leaks at all. GreenPath has removed vent records from the leak data via algorithm and human classification of the top leaks but all 3,082 leaks in the data set have been individually inspected to determine if they are leaks or vents. Additionally, there are 2,664 leak points with an emission rate of zero, a record of a leak resolved (based on the most common data platform).

The inconsistencies in the data create challenges in drawing inferences from the data, combined with the lack of time series data on facilities to show the effectiveness of leak detection and repair programs.

Table 16: Number of facilities by number of consecutive years surveyed

Number of years surveyed	Number of facilities
5	11
4	45
3	50
2	89
1	689

A visual representation of the count of leaks and sum of volumes at the 11 most-inspected facilities shows no discernable pattern, as can be seen in the figure below. Among sites inspected four times there is a more discernable pattern. One possible cause may be that in the most inspected sites, there are sometimes multiple inspections per year and thus two inspections are added together to make one inspection for the year, and currently those inspections are not broken out. A revised Python script should be able to parse out multiple inspections per year.

One potential source of the variability is the introduction of the FLIR GF320 in 2012-13. The GF320 makes the detection of leaks much more user friendly relative to the original Gasfind IR camera, and would offer a partial explanation for the spike in leaks found in 2012-13 at the most inspected sites.

The following visuals show results for the most consistently-inspected facilities, each with five years of recorded inspections.

Figure 11: Top 11 most inspected sites - by leak rate (CFM) (five inspections per year)

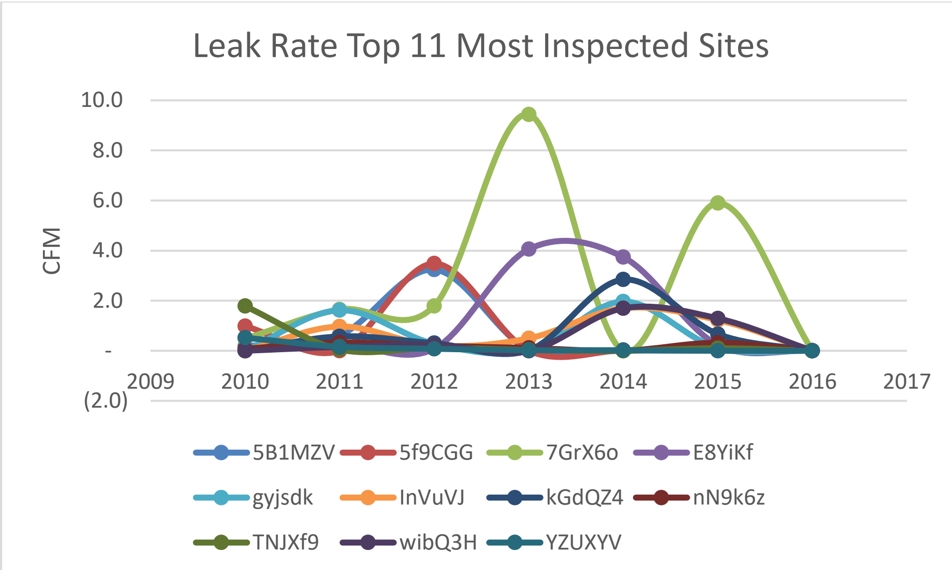
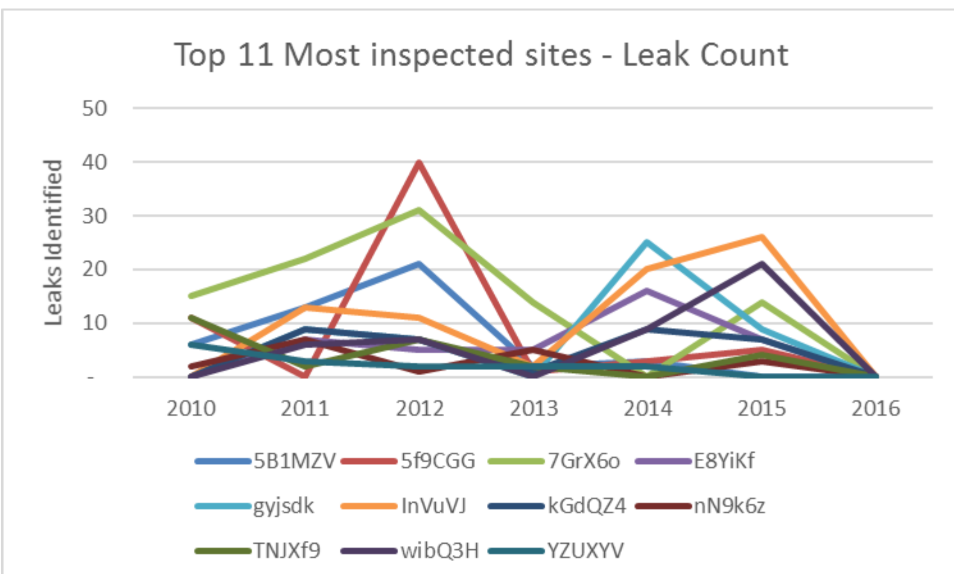


Figure 12: Top 11 most inspected sites - by leak count



Site 7GrX6o shows a high leak count in 2012, but a high leak rate and low leak count in 2013. The reason for this disconnect between leak volume and leak count between the two surveys is a 6CFM leak from a connector.

Figure 13: Sites with four Inspections - leak count

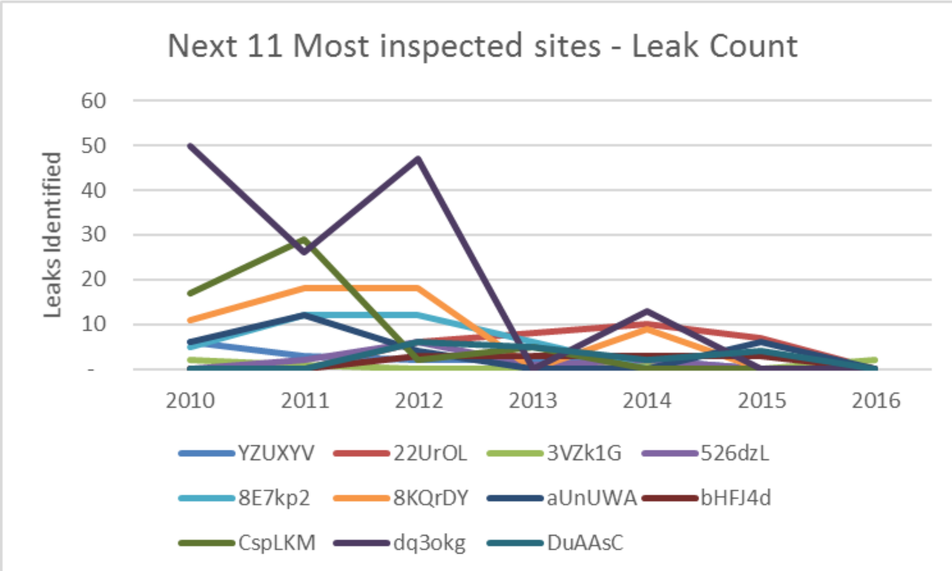
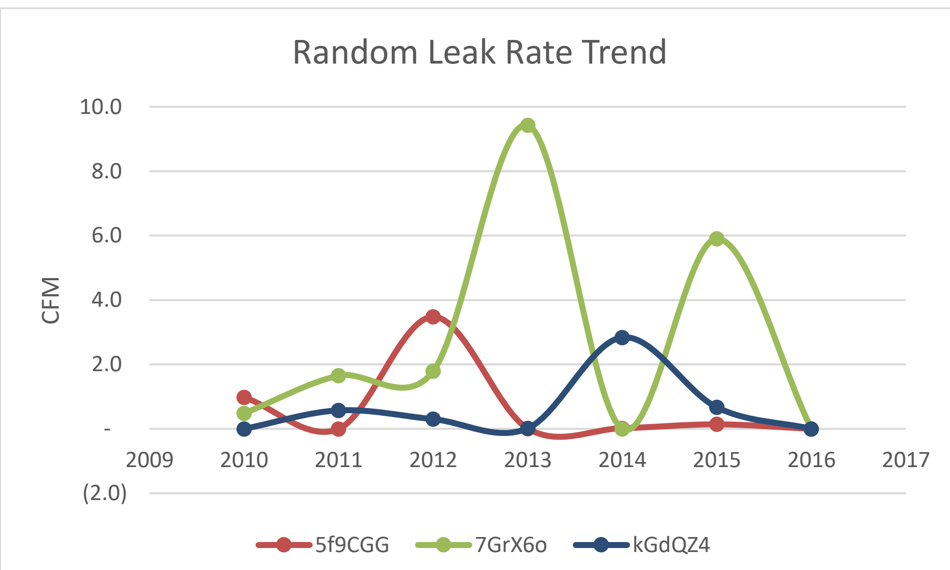


Figure 14: Oscillating leak rates



Further analysis is required to dig through the data to understand why random patterns such as those shown in Figure 11 exist, as well as understand what drives downward trends such as Figure 12 or 13, or cases where low emissions suddenly spike.

Figure 15: Declining leak rate trend

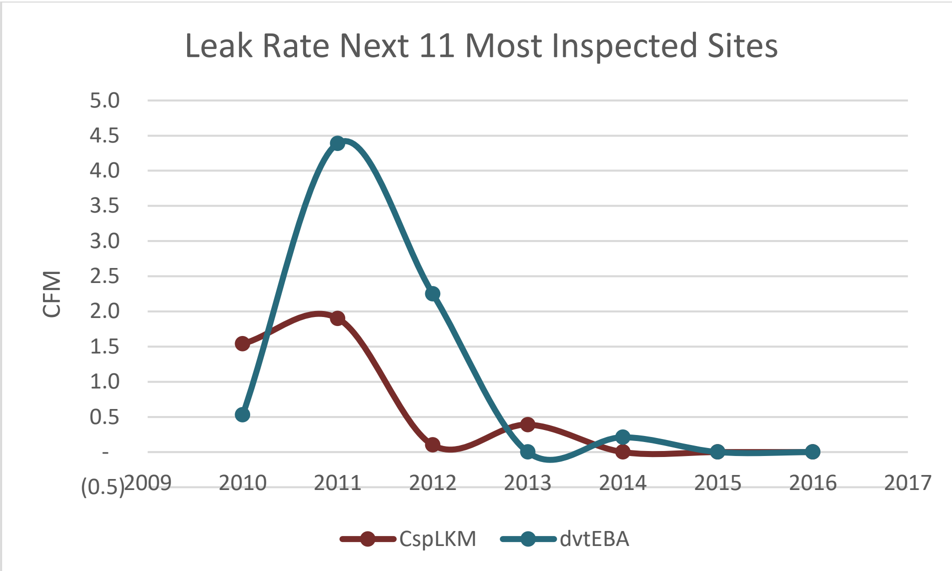
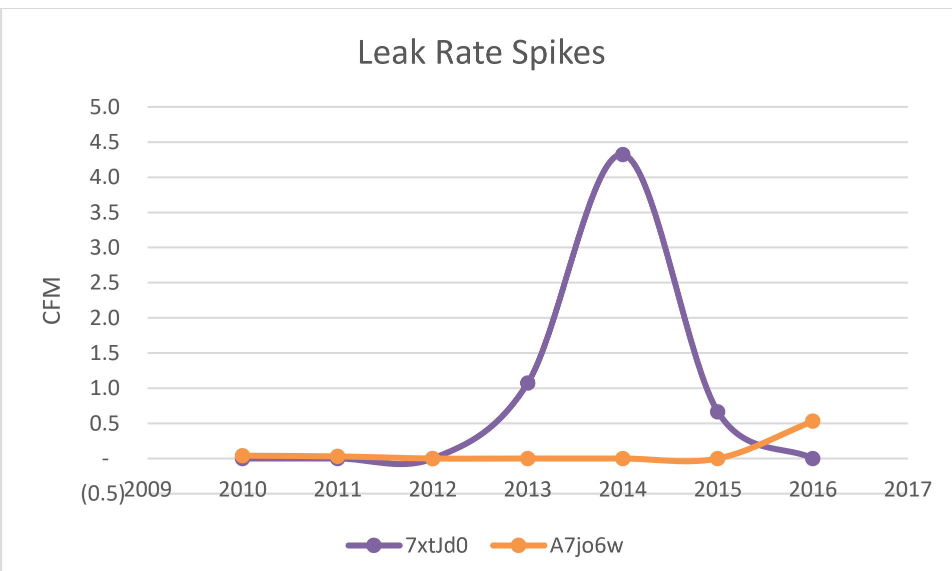


Figure 16: Facilities with leak rate spike



SUPER-EMITTER FINDINGS

In analyzing the data, 37 leaks greater than 6 CFM were identified which could be classified as super-emitters. One very large leak was detected (estimated at 150CFM via calibrated volume bag) and resolved within two weeks.

An issue within the data is the lack of consistency whether an emission rate was quantified or estimated. By looking at the individual records it can sometimes be determine if a leak value was estimated or quantified (via high flow sampler or other methods), but most data sets did not include a field for "estimated or measured". Further analysis is required to determine how many of these 37 super-emitter leaks were resolved; these account for approximately 23% of the total leak volume in the survey. Then, removing the one 150 CFM leak, these "super-emitters" account for 12% of total leak volume.

Only 14 facilities in the survey would meet the definition of a facility super-emitter (total site emissions greater than 30 CFM). Of those 14 facilities, only three cases included emissions driven by leaks, with the remainder driven by tank venting emissions.

It is important to note that, currently, LDAR teams do not deploy with measurement tools capable of quantifying emissions from tank vents; therefore, these estimates are largely based on operator "best guesses" with significant variability around the accuracy in relation to tank vents.

A further issue to raise is the tracking of repairs, which is inconsistent within the data set obtained from producers. Only a small subset of the data was recorded when a leak was resolved, and those repair records are not well integrated within the data. Further analysis would be required to determine the fate of super-emitter leaks, which would require being able to link leak detection database systems with operational and maintenance logs, or follow-up surveys on the same facility to ensure that super-emitter leaks have been resolved.

CONCLUSION AND RECOMMENDATIONS

Historical LDAR practice in Canada is currently highly variable. The variation in data collection, maintenance and analysis practices was not known prior to undertaking the study. Over 80% of the hours allocated to this project related to processing data to the point where meaningful analysis could be made. The tools developed to process the data will allow for additional information to be added at a relatively low time and cost commitment, but building the tools to create a viable data set were very costly and time-consuming.

The largest criticism of LDAR in Canada is the lack of consistency in data collection, not just with regards to leak detection, but with attention to the tracking of repairs. The existing CAPP *Best Management Practice* has provided flexibility to operators to design their own internal FEMP, this in turn, results in each operator having different data collection methods creating challenges with inconsistencies in data collection, quality assurance and quality control, and most notably the tracking of repairs.

As Canadian and American legislators strive to achieve a 45% reduction in methane from the upstream oil and gas sector, enhancements to leak detection and repair practice will be part of the regulatory framework. For credit for progress towards this goal to be claimed by the upstream oil and gas industry, better data management, guidance, and methodologies for leak detection and repair must be developed. In Appendix E of this report, GreenPath has developed a spreadsheet-based tool for data collection which will standardize further leak detection data collection efforts and make updates to this report less onerous.

The development of a comprehensive, credible and consistent data set will assist with the ultimate objective of minimizing cost to industry while maximizing methane reductions.