2020-10-13

Evaluating Technologies and Methods for Measuring Methane Emissions from the Upstream Oil and Gas Sector

Fox, Thomas Arcadius Oram

UNIVERSITY OF CALGARY

Evaluating Technologies and Methods for Measuring Methane Emissions from the Upstream Oil and Gas Sector

by

Thomas Arcadius Oram Fox

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN GEOGRAPHY

CALGARY, ALBERTA

OCTOBER, 2020

Abstract

Methane, a potent greenhouse gas, is commonly emitted during the production, processing, transmission, and storage of oil and natural gas (O&G). The O&G industry is the leading source of anthropogenic methane in Canada and the US, but methane measurement and mitigation strategies remain underdeveloped. In recent years, an increasing number of O&G producing jurisdictions have introduced regulations mandating methane leak detection and repair (LDAR) programs in order to reduce fugitive emissions. Meanwhile, innovation in methane measurement has exploded, with companies emerging that promise to reduce methane emissions using drones, aircraft, satellites, fixed installations, handheld instruments, and other vehicle systems. These new solutions are not well understood, and how they might contribute to reducing methane emissions is unclear. This thesis seeks to improve understanding of emerging methane-sensing technology performance and participation in the O&G industry. Specifically, it seeks to reveal whether current and emerging technologies are technically capable of reducing emissions, can meet regulatory requirements for approval, and offer cost savings relative to established methods. This thesis presents four chapters of research with the following main results: (1) Screening is a way for mobile technologies to rapidly search for large leaks. Most emerging technologies and methods use screening and can detect methane in some capacity, but much more testing is needed to understand performance metrics, precise limitations, how to direct follow-up, and mitigation potential; (2) Policies are evolving to enable adoption of new systems, but careful work will be required to properly evaluate suitability through controlled testing, simulation modeling, and piloting; (3) LDAR-Sim and similar tools can support the development of LDAR programs with new technologies but modeling results are highly sensitive to technology performance assumptions, empirical inputs, and environmental conditions; (4) Screening technologies that require follow up may struggle to be cost-effective due to high quantification errors and the confounding presence of vented and combustion emissions at most facilities. Impressive progress has been made in developing, enabling, and implementing new LDAR technologies, but regulators, industry, and researchers should continue to work together to ensure credible and defensive emissions reductions are achieved through implementation.

Preface: Publications and Co-Authorship

Chapter 2

Thomas A. Fox,[1] Thomas E. Barchyn,[1] David Risk,[2] Arvind P. Ravikumar,[3] Chris H. Hugenholtz[1]

[1] Department of Geography, University of Calgary, Calgary, T2N 1N4, Canada

[1] Department of Earth Sciences, St. Francis Xavier University, Antigonish, B2G 2W5, Canada

[3] Harrisburg University of Science and Technology, Harrisburg, PA 17101 USA

All authors contributed to the writing and editing of text. TAF and APR produced the figures. The authors declare that they have no conflict of interest. The authors acknowledge assistance from Siobhan Semadeni, Michael Daciw, Roopa Ganapathy, Stephane Germain, Steve Deiker, Ventus Geospatial, Dirk Richter, Natural Resources Canada, and Petroleum Technology Alliance of Canada.

Chapter 3

Thomas A. Fox,[1] Arvind P. Ravikumar,[2] Chris H. Hugenholtz,[1] Daniel Zimmerle,[3] Thomas E. Barchyn,[1] Matthew R. Johnson,[4] David Lyon,[5] Tim Taylor[6]

[1] Centre for Smart Emissions Sensing Technologies, University of Calgary, 2500 University Drive NW, Calgary, AB, T2N 1N4, Canada

[2] Harrisburg University of Science and Technology, 326 Market St, Harrisburg, PA 17101, USA

[3] Energy Institute, Colorado State University, 400 University Ave, Fort Collins, CO 80523, USA

[4] Energy & Emissions Research Lab, Mechanical and Aerospace Engineering, Carleton University, 1125 Colonel By Dr, Ottawa, ON, K1S 5B6, Canada

[5] Environmental Defense Fund, 301 Congress Ave #1300, Austin, TX 78701, USA

[6] Colorado Department of Public Health and Environment, 4300 Cherry Creek S Dr, Denver, CO 80246, USA

Workshop was planned and organized by TAF, APR, and CHH. All authors contributed to framework conception and development. TAF drafted the initial manuscript and all authors contributed to editing. The authors declare that they have no conflict of interest. The authors acknowledge assistance from Environment and Climate Change Canada for initial funding, Karen Marsh for edits, and workshop participants for contributions and review.

Chapter 4

Thomas A. Fox,[1] Mozhou Gao,[1] Thomas E. Barchyn,[1] Yorwearth L. Jamin,[2] Chris H. Hugenholtz[1]

[1] Centre for Smart Emissions Sensing Technologies, University of Calgary, 2500 University Drive NW, Calgary, AB, T2N 1N4, Canada

[2] Clearstone Engineering Ltd., 900 6 Ave SW #700, Calgary, AB, T2P 3K2, Canada

TAF, MG, TEB, and CHH contributed to project conception and design. TAF, MG, and TEB designed and wrote the model code. TAF conducted the case study, including parameterization, simulation, and analysis. TAF designed and conducted the sensitivity analysis and drafted the manuscript. All authors contributed to writing, review and approval of the final manuscript.

Chapter 5

Thomas A. Fox,[1] Chris H. Hugenholtz,[1] Thomas E. Barchyn,[1] Tyler Gough,[1] Mozhou Gao[1]

[1] Centre for Smart Emissions Sensing Technologies, University of Calgary, 2500 University Drive NW, Calgary, AB, T2N 1N4, Canada

All authors contributed to study conception and manuscript review and editing. TAF led study design, simulations, data analysis, production of figures, and writing.

Acknowledgements

Thank you to my supervisor Dr. Chris Hugenholtz for encouraging and believing me in ever since we met more than 10 years ago when I was an undergrad at the University of Lethbridge. Over the years, I have learned a great deal from Chris, who inspires me to bring more enthusiasm to my work, to focus on solving real-world problems, and to dive in head-first when learning something new.

Thank you to all my lab mates for their friendship and support - particularly Elena, Paul, and Maja who welcomed me in the beginning and remained my companions throughout. Elena deserves special recognition for her thoughtful and caring disposition, exemplified not least by the *dozens* of birthday cards she has unfailingly circulated (inviting others to sign their names). Thank you to Mo and Tom for sharing their technical knowledge that enabled me to learn how to code - something I never thought I could achieve. Thank you also to my friends and colleagues in the methane space, especially Jessica and Brenna, who have patiently taught me a great deal about how different the world is outside of academia.

Thank you to my supervisory committee, Drs. Brent Else and Ke Du, and to Drs. Gwendolyn Blue and Arvind Ravikumar for helping me through candidacy. Thank you to all the faculty and admin in U of C Geography for working to help students thrive and be happy. Thank you to NSERC and AITF (but ultimately taxpayers) for the scholarships that gave me the privilege of not having to worry about money.

Most importantly, I feel immense gratitude for my partner Steph, my mom and dad, my brother and sister, and my close friends. No amount of reflecting on the unconditional and unwavering support and love I have received can leave me with the words to adequately describe how fortunate I feel. While I have worked hard, the most important factor in getting to this point has been having people who have cared for me, supported me, and have been invested in my success. Steph deserves an award for tolerating me during the most stressful times, for helping me to see the good in people, and for injecting bursts of joy into long days behind a screen. I truly appreciate your silliness and care, despite stubbornly resisting when you try to drag me out of my workaholic trance and into the blinding sun.

Table of Contents

Abstract ii
Preface: Publications and Co-Authorship iii
Acknowledgements v
List of Figures ix
List of Tables xii
List of Abbreviations xiii
Chapter 1 Introduction 14
1.1 Methane emissions from oil and gas 14
1.2 Regulatory landscape 16
1.3 Knowledge gaps 18
1.3.1 Knowledge gap 1: Current technology and methods 18
1.3.2 Knowledge gap 2: Integrating alternative methods 18
1.3.3 Knowledge gap 3: Effectiveness of multi-visit screening 19
1.4 Thesis overview 19
1.4.1 Chapter 2 19
1.4.2 Chapter 3 20
1.4.3 Chapter 4 20
1.4.4 Chapter 5 21
Chapter 2 A review of close-range and screening technologies for mitigating fugitive methane emissions in upstream oil and gas 22
2.1 Abstract 22
2.2 Introduction 22
2.3 Review methods and framework 25
2.4 Technology overview 27
2.4.1 Measuring methane 27
2.4.2 Technology classes 28
2.5 Technology review 30
2.5.1 Handheld instruments 30
2.5.2 Fixed sensors 33
2.5.3 Mobile ground labs 34
2.5.4 Piloted aircraft 36
2.5.5 Unmanned aerial vehicles 39

2.5.6 Satellites 40
2.6 The future of LDAR 42
2.6.1 Integrating screening technologies 42
2.6.2 Beyond technology integration 45
2.6.3 Regulatory and economic considerations 46
2.7 Conclusion 47
Chapter 3 A methane emissions reduction equivalence framework for alternative leak detection and repair programs 48
3.1 Abstract 48
3.2 Introduction 48
3.3 Framework development 49
3.4 Definitions 50
3.5 Equivalence framework 51
3.5.1 Stage 1: Method identification 53
3.5.2 Stage 2: Controlled testing 53
3.5.3 Stage 3: Simulation modeling 54
3.5.4 Stage 4: Field trials 55
3.5.5 Stage 5: Full approval 55
3.5.6 Exceptions 56
3.6 Challenges 56
3.6.1 Controlled testing 56
3.6.2 Modeling 57
3.6.3 Scale and source disambiguation 57
3.6.4 Human factors 58
3.6.5 Logistics 58
3.7 Conclusions 58
Chapter 4 An agent-based model for estimating emissions reduction equivalence among leak detection and repair programs 60
4.1 Abstract 60
4.2 Introduction 60
4.3 Model description 63
4.3.1 Program definition 63
4.3.2 Initialization 65
4.3.3 Updating 66

4.3.4 Finalization ... 68
4.4 Case study ... 70
4.5 Model confidence ... 73
Chapter 5 Will new mobile screening technologies enable cost-effective methane leak detection? ... 79
5.1 Abstract ... 79
5.2 Introduction ... 79
5.3 Methods ... 82
5.3.1 Triaging procedures ... 84
5.3.2 Equivalence scenario modeling ... 85
5.3.3 Cost-effectiveness modeling ... 87
5.4 Results ... 89
5.4.1 Equivalence scenarios ... 89
5.4.2 Cost-effectiveness ... 93
5.5 Discussion ... 94
5.6 Conclusion ... 97
Chapter 6 Conclusion ... 98
6.1 Summary of contributions ... 98
6.2 Implications and future work ... 101
6.3 Final thoughts ... 104
Chapter 7 Appendix A: Supplementary information for Chapter 4 ... 105
7.1 Operator detection ... 105
7.2 Case study parameterizations ... 107
7.3 Sensitivity analysis ... 109
7.4 Supporting tables and figures ... 113
Chapter 8 Appendix B: Supplementary information for Chapter 5 ... 119
Chapter 9 Literature cited ... 120

List of Figures

Figure 2.1 Kernel density map showing spatial variability of well density in Alberta and Northeastern British Columbia, Canada. Important oil and gas municipalities include 1) Calgary; 2) Cold Lake; 3) Edmonton; 4) Fort McMurray; 5) Grande Prairie; 6) Medicine Hat; and 7) Red Deer. Inset histogram represents the distance from each well to the nearest of 1010 population centres (μ = 38 km, range = 0-290 km). Well data was acquired from the Alberta Energy Regulator in June 2017. .... **24**

Figure 2.2 Technology classes categorized based on the spatial and temporal extent of coverage. Coloured dots represent suitability for measurement motivations 1-4. Dots without black borders either show promise or may be useful in a limited capacity. **29**

Figure 2.3 Examples of data output scale and form: a) FLIR GF320 image of an emitting pipe; b) Methane flux wind rose of a fixed sensor (provided by Quanta3); c) MGL concentration map near a large cattle feedlot; d) Piloted aircraft retrieval using imaging spectroscopy (provided by Kairos Aerospace); e) Column-averaged methane enhancements from a flooded dam (provided by GHGSat); and f) fixed-wing UAV concentration map of a controlled release. White arrow indicates wind direction. **32**

Figure 2.4 Monthly average NASA cloud fraction (Terra/MODIS, 0.1 degree) of North America for January and July 2017. Cloud cover can limit the application of aircraft- and satellite-based imaging spectrometers. **38**

Figure 3.1 Number of workshop participants by stakeholder group. **50**

Figure 3.2 Overview of the equivalence framework. Stages are separated by objectives, whether informally (fine line) or by regulatory approval (dashed line). An adaptive feedback process connects each of the five stages to enable sharing of knowledge and evolution of the demonstration process. **52**

Figure 4.1 Overview of LDAR-Sim. Green text is used for screening methods and orange text for close-range methods. Red arrows represent ‘no’, green arrows are ‘yes’, and grey arrows are mandatory. **62**

Figure 4.2 Fugitive emissions observed under each program. Lines show daily emissions averaged across 25 simulations ($\pm 2\sigma$). A and C show average site-level emissions, while B and D show candidate programs differenced against the regulatory program ($P_X - P_{ref}$). In B and D, equivalence occurs at y = 0. .... **73**

Figure 4.3 Sensitivity analysis results. Darker colours indicate more sensitive inputs; see Section 7.3 for derivation of sensitivity index S. Outputs are: (A) median number of active leaks, (B) median number of days leaks remain active before discovered, (C) mean daily site emissions, (D) cumulative number of repaired leaks. For each output, column 1 calculates S using absolute differences between programs ($MGL - OGI$), while column 2 uses a ratio (MGL/OGI). Values of S near zero suggest that the input parameter has little influence over the equivalence metric, though don't necessarily indicate that the program is unaffected by the input. **75**

Figure 4.4 Differences (A; MGL - OGI) and ratios (B; MGL/OGI) of fugitive emissions as a function of LPR. Each dot compares a close-range and a screening program that share global parameters. Purple and green lines show operator presence and absence, respectively (i.e. NRR turned on or off). Dashed lines show presence and solid lines show absence of vented emissions. The dashed vertical line is the mean LPR from the SA sampling distribution (0.026), established using Alberta data. The solid vertical line is the LPR value used in FEAST and the shaded region is the LPR SA range evaluated in Kemp et al. (2016) **78**

Figure 5.1 Ensemble of simulations across a range of dynamic follow-up targets and annual surveys. As the dynamic target increases, more facilities receive follow-up, more leaks are found, and fugitive emissions are lower. Increasing survey frequency similarly results in lower fugitive emissions as infrequent large emissions are found and repaired sooner. Intersections between regulatory OGI programs (black horizontal lines) and screening programs (colour lines) denote equivalence. For example, to be equivalent with triannual OGI, a screening program that screens all facilities 9 times per year must send OGI follow-up to the top 12% of highest-emitting facilities. Shaded regions are two standard deviations ensemble variability in emissions rate. This example shows simulations under D1, E1, and without design emissions (the most favourable situation for screening technologies to be economically viable). Fugitive emission rates are averaged over 5-year simulations. **90**

Figure 5.2 Catalogue of equivalence scenarios under various conditions: (A) Dynamic targets assuming leak-rate distribution D1 (more skewed), (B) Dynamic targets assuming leak-rate distribution D2 (less skewed), and (C) Static targets assuming D1. Each point represents an equivalence scenario with either annual (orange) or triannual (blue) conventional OGI programs. Solid lines optimistically assume that all emissions are fugitives, while dashes lines assume the presence of design emissions. Quantification errors increase from E1 (perfect facility-scale quantification) to E3 (order or magnitude error) and are explained in the main text. Note the different interpretations of dynamic targets (τ; proportion of top-emitting facilities requiring follow-up) and static targets (proportion of the fugitive leak-size distribution to target). **92**

Figure 5.3 Maximum screening program costs (C_screen) under a range of equivalence scenarios. Estimates are made assuming OGI survey costs of USD $250/facility (lower limit of each line) and USD $450/facility (upper limit of each line). Blue lines show required screening costs when screening program target costs are equal to those of the regulatory program ($\varphi = 0$). Orange lines show required screening costs for screening programs that offer a 30% cost reduction relative to the OGI-only regulatory standard ($\varphi = 0.3$). The best-case scenario for MVL assumes perfect screening quantification (E1), leak rates from D1, and no design emissions. The worst-case scenario assumes E3, D2, and presence of design emissions. Negative C_screen values indicate that screening technology providers must pay to perform services., a situation unlikely to exist in practice. **94**
Figure 7.1 Illustration of the three types of redundancies tracked in LDAR-Sim. **114**
Figure 7.2 The spatial distribution of OGI suitability in Alberta for different deployment thresholds. On the left, crews are deployed if warmer than -20 °C and winds are below 20 $m \cdot s^{-1}$. On the right, stricter thresholds are used, resulting in fewer deployment days. **115**
Figure 7.3 Detection probability curves for six possible $pmax$ values using the Fort Worth Air Quality Study. For example, the gold line ($pmax$ = 0.05) assumes operators have a 5% chance of detecting the largest leak in the dataset in NRR calculations. **115**
Figure 7.4 Empirical frequency distribution of leak production rates in Alberta. The mean (vertical black line) is 0.026. The red line is the LPR value used in FEAST and in this LDAR-Sim demonstration. The shaded region is the LPR sensitivity analysis range evaluated in Kemp et al. (2016). **116**
Figure 7.5 SA results for a single OGI program. Rows are inputs and columns are outputs: (A) number of active leaks, (B) number of days leaks are active, (C) daily emissions, (D) missed leaks, (E) proportion of sites available for LDAR, (F) repaired leaks, and (G) surveys conducted. **117**
Figure 7.6 Sampling distributions used in the SA (N = 50 000). **118**
Figure 8.1 Equivalence scenarios (D1; E = 1; no design emissions) for leak production rates one order of magnitude below (0.00065, left) and above (0.065, right) the LPR used for this study (0.0065, middle). All simulations performed without null repair. **119**

List of Tables

Table 2.1 Comparison of leak detection technologies and methods. **31**

Table 4.1 Parameters used for the case study. These programs are hypothetical and based on fabricated detection methods and survey requirements for the sake of demonstration. **72**

Table 5.1 Static targets and corresponding static thresholds used in this study. Two leak-size distributions are used; D1: Clearstone Engineering Ltd. (2018), and D2: Ravikumar et al. (2020). **85**

Table 7.1 LDAR-Sim output data by day, site, and leak... **113**

Table 7.2 Case study results.. **113**

Table 7.3 Approximate number of active leaks and daily emissions for 500 sites under a range of $pmax$ values (simulations conducted without LDAR). Note that if the largest leaks have only a 5% chance of being detected by operators during weekly visits (gold line in Figure 7.3), system-wide emissions fall by >50%. **114**

List of Abbreviations

alt-LDAR	Alternative LDAR
AVO	Audio, Visual, and Olfactory
AWP	Alternative Work Practice
CMP	Comprehensive Monitoring Program
DIAL	Differential Absorption LiDAR
ECCC	Environment and Climate Change Canada
EPA	Environmental Protection Agency
FEAST	Fugitive Emissions Abatement Simulation Toolkit
GHG	Greenhouse Gas
IPCC	Intergovernmental Panel on Climate Change
LDAR	Leak Detection and Repair
LiDAR	Light Detection and Ranging
LPR	Leak Production Rate
LSA	Lowest Safe Altitude
MGL	Mobile Ground Lab
NG	Natural Gas
NRR	Null Repair Rate
O&G	Oil and Gas
OGI	Optical Gas Imaging (camera)
OTM	Other Test Method
OVA	Organic Vapour Analyzer
reg-LDAR	Regulatory LDAR
SA	Sensitivity Analysis
SWIR	Short-Wave Infrared
UAV	Unmanned Aerial Vehicle
US	United States

Chapter 1

Introduction

1.1 Methane emissions from oil and gas

Global atmospheric methane concentrations have more than doubled since the start of the 19th century due to anthropogenic disruption of natural source-sink dynamics (Stocker, 2014). Methane is a potent greenhouse gas (GHG); a molecule of methane has a global warming potential 25 times that of carbon dioxide over 100 years (Boucher et al., 2009). Although atmospheric methane is short-lived (approximately 10 years compared to hundreds to thousands of years for carbon dioxide), mean global concentrations have increased from around 720 ppb in 1750 to over 1850 ppb today, accelerating since the early 2000s (Etheridge et al., 1998; Nisbet et al., 2019). Common anthropogenic sources include the oil and gas (O&G) sector, paddy fields, biomass burning, livestock, landfills, and wastewater (National Academies of Sciences and Medicine, 2018).

The O&G industry is the largest source of anthropogenic methane emissions in both Canada and the United States (ECCC, 2016; EPA, 2019). Natural gas (NG), which consists primarily of methane, can be emitted in a variety of ways. In North America, most regulations distinguish between vented and fugitive emissions. Vents are intentional, occur by design, and often account for the majority of emissions at individual facilities (Allen et al., 2013). Examples of venting include instrument gas used to power pneumatic devices and incomplete combustion of NG emitted in the exhaust of compressor engines. Vented emissions tend to be higher in facilities that produce oil because NG is not the target commodity (Lyman et al., 2019). Reducing vented emissions generally requires changing the processes that cause them. Examples include using low-bleed pneumatics, installing vapour recovery units on liquid storage tanks, or flaring excess methane (i.e., combusting it into carbon dioxide).

Fugitive emissions, sometimes called leaks, are unintentional emissions from broken, malfunctioning, or misused equipment. Because methane is often invisible and odourless in upstream settings, finding fugitive sources can be challenging and generally requires the use of specialized instruments (Ravikumar et al., 2018). To reduce fugitive emissions, regulators in North America increasingly require that O&G companies implement a leak detection and repair

(LDAR) program (otherwise known as a fugitive emissions management program, or FEMP). Until recently, these programs have generally consisted of targeted inspections of each potentially leaking component using organic vapour analyzers (OVA) or optical gas imaging (OGI) cameras. However, the industry is changing rapidly, and 'alternative' methods such as vehicles, aircraft, drones, satellites, or fixed sensors are gaining popularity (Fox et al., 2019a).

Until recently, few concerted efforts sought to measure methane emissions from O&G operations in North America. That changed in 2012, when the Environmental Defense Fund (EDF) launched a six-year coordinated field campaign to better understand the methane sources and their contribution to federal emissions. Contributing to the campaign were 140 academics and experts representing 40 institutions and 50 companies; most work took place in the Barnett Shale, Texas, USA. Overall, the mega-study suggested that the US government was underreporting methane emissions from the NG supply chain by approximately 60% (Littlefield et al., 2017; Alvarez et al., 2018).

During and following the EDF study, a great deal of related research was conducted. The majority of this research has measured emissions in the field using different approaches, seeking to understand how emissions vary by basin, production type, facility, activity type, and over time (e.g., Johnson et al. 2017; Omara et al. 2018; Peischl et al. 2018; Duren et al. 2019). Much of the work to date has revealed and attempted to reconcile apparently conflicting 'top-down' (i.e., aggregate-scale emissions of an entire region or facility) measurements with much lower 'bottom-up' measurements (i.e., component level measurements averaged and extrapolated across larger scales) (Zavala-Araiza et al., 2015; Vaughn et al., 2018).

The historical discrepancy between top-down and bottom-up emissions exemplifies our limited understanding of how much, where, and when O&G systems emit methane. Several hypotheses have been proposed to explain the discrepancy. A popular hypothesis is that bottom-up inventories often fail to account for rare but large emission sources (Zavala-Araiza et al., 2015, 2017). Distributions of methane emission rates from O&G systems are characteristically heavy-tailed, with 5% of sources commonly accounting for over 50% of emissions (Brandt et al., 2016). The largest sources occur infrequently and exhibit anomalous emission rates, meaning that they are unlikely to be considered when developing the emissions factors that make up inventories. A second hypothesis is that certain high-impact events, such as blowdowns or tank venting, which are difficult or impossible to quantify at the component scale, are inadequately represented in bottom-up inventories (Vaughn et al., 2018). Both of these hypotheses suggest that bottom-up measurements provide an incomplete picture. This is worrisome, as greenhouse gas tracking, and mitigation efforts in Canada and the US are grounded in bottom-up inventories.

Despite this emerging body of work, our understanding of fugitive emissions is largely descriptive, and very little is known about what causes equipment to leak. As a result, most leak detection programs are deployed blindly and comprehensively, following the implicit assumption that all components are equally likely to leak. Fugitive emissions simulation models generate leaks following the same assumption (Kemp et al., 2016). Given the sparse distribution typical of O&G systems, exhaustive component-scale inspections tend to be expensive, especially if multiple surveys are required each year. For this reason, interest is growing in alternative technologies and methods able to accelerate the identification of fugitive sources – especially the largest ones.

1.2 Regulatory landscape

Several incentives to prevent leakage exist that are independent of government intervention. Fugitive emissions are lost product in the NG industry, and mitigation can be considered a form of product-loss prevention. Methane emissions may also pose an explosive hazard and often have associated air pollutants that may be harmful to human health (Garcia-Gonzales et al., 2019). Greenhouse gas emissions are also becoming a concern for shareholders, who are increasingly predicting future financial performance using environmental and social corporate governance. However, these natural incentives are generally deemed insufficient to achieve meaningful methane emissions mitigation, and an increasing number of O&G producing jurisdictions in North America are now mandating LDAR.

In North America, commitments by Canada, the US (since retracted), and Mexico to reduce methane emission from the O&G sector by 40-45% by 2025 have led to a flurry of new regulations. In January 2020, new federal regulations took effect in Canada, developed and enforced by Environment and Climate Change Canada (ECCC, 2018). The ECCC regulations target vented emissions, fugitive emissions, and methane emissions from incomplete combustion (e.g. from flare stacks). To reduce fugitive emissions, ECCC requires any O&G facility larger than a single well to be surveyed three times per year using a handheld organic vapour analyzer or an OGI camera. The precise number of facilities affected by the regulations is unclear because inclusion depends on what types of equipment are present at each site. However, a conservative estimate of 50 000 facilities out of an approximate 300 000 active wells in Alberta would require 150 000 facility inspections. At $500 per inspection, these regulations will cost industry upwards of $75 million annually.

Concurrently in January 2020, an independent set of regulations developed by the Alberta Energy Regulator took effect (AER, 2018). The AER regulations only target certain facility subtype codes and require only annual inspections for most facilities alongside triannual inspections for larger facilities. In Alberta, Johnson and Tyner (2020) catalogued 19,607 facilities subject to annual LDAR and 2,308 sites subject to triannual LDAR. Negotiations are currently underway between Alberta and ECCC to find equivalency between the two sets of regulations so that the AER requirements can supersede the more onerous ECCC requirements. These negotiations are not public, and it remains unknown if or when a settlement will be made.

Both AER and ECCC regulations distinguish between ‘regulatory’ and ‘alternative’ LDAR programs. In regulatory LDAR, compliance is precisely defined: inspectors must use OVAs or OGIs, abide by clear work practices (e.g. measurement time and distance), conduct surveys at a set frequency separated with minimum intervals, and so on. In alternative LDAR, asset-holders may devise a custom LDAR program comprised of their choice of technology, work practice, and program specifications (e.g. survey frequency). As an example, a company could use satellite measurements of point-source methane emissions to inform drone-based LDAR. In Alberta, alternative programs must be approved by the AER. Federally, ECCC must be informed, but need not approve prior to implementation.

The option to conduct alternative LDAR was introduced for several reasons. First, asset-holders were concerned about the cost of compliance of regulatory LDAR. Second, many new methane-sensing technologies have blossomed during the past decade, including mobile ground labs (MGL) (Caulton et al., 2018), unmanned aerial vehicles (UAVs) (Barchyn et al., 2019), aircraft systems (Conley et al., 2016), satellites (Varon et al., 2018), and fixed sensors (Alden et al., 2017). While the suitability of these options for LDAR remains uncertain, it may be unwise to close the door entirely on their use, especially given how long it takes to develop and implement new regulations. Third, considerable variability exists in facility type, density, road access, climate, and geology, among other variables. Different solutions may therefore be expected to perform better in different contexts. Fourth, granting flexibility in technologies and methods to be used may improve competition among solution providers and lead to innovation.

Similar laws are being developed and enforced in British Columbia, Saskatchewan, and across the US. However, ECCC and AER regulations are leading the world in terms of embracing alternative LDAR. For this reason, and because the majority of Canada's O&G activities are in Alberta, my thesis is loosely focused on the Alberta context.

1.3 Knowledge gaps

In recent years, alternative methods have been used with varying success to detect, locate, and quantify methane emissions from upstream oil and gas (Schwietzke et al., 2018; Ravikumar et al., 2019). However, these methods are new, and each faces a suite of operational, environmental, and physical limitations, making them effective under only specific conditions. In general, all of the available methane-sensing technologies, including traditional sensors, remain poorly understood with regards to mitigation effectiveness under operational (i.e. field) conditions. Regulators, academics, and the O&G industry are excited to begin integrating emerging methods into LDAR programs, but several knowledge gaps remain that should be addressed before new methods are adopted.

1.3.1 Knowledge gap 1: Current technology and methods

What technologies and methods currently exist, what are they capable of doing, and what are their limitations? Few market-ready methane-sensing technologies exist, and even fewer are approved for use by regulators. Until 2018, only handheld devices such as optical gas imaging (OGI) cameras and organic vapour analyzers (OVAs) were approved by regulators for conducting leak detection and repair (LDAR). While relatively precise, these instruments tend to be expensive and labour-intensive, as distances-to-source must generally remain below 10 m. Recently, interest has grown in using mobile ground labs (MGLs), unmanned aerial vehicles (UAVs), aircraft, satellites, and fixed sensors for monitoring methane emissions in O&G. However, the performance of all available methane-sensing technologies, including traditional sensors, remains poorly understood under operational (i.e. field) conditions. This knowledge gap is addressed in a literature review presented in Chapter 2.

1.3.2 Knowledge gap 2: Integrating alternative methods

Most current regulations that enable the deployment of alternative technologies, methods, and programs require that producers demonstrate that alternatives lead to equivalent mitigation outcomes when compared to a regulatory standard (AER, 2018; ECCC, 2018). However, regulations tend to be vague when it comes to providing instructions on how to demonstrate equivalence. As a result, producers, solution providers, and other stakeholders are left asking: 'How does one demonstrate equivalence?' The lack of specificity on behalf of regulators may be intentional; showing that two technologies, methods, or programs achieve equivalent emissions mitigation is challenging. First, there is little agreement over how well regulatory methods (e.g. OGI) perform, meaning there exists no clear goalpost for demonstrating equivalence. Second, most emerging technologies have not been tested using controlled releases to evaluate performance. Furthermore, no standardized testing protocol exists, and evaluating technologies and methods under a representative range of environmental conditions, basins, and facility types is a daunting task. Third, no widely accepted simulation framework exists for evaluating LDAR programs. This knowledge gap is addressed in Chapter 3 and Chapter 4.

1.3.3 Knowledge gap 3: Effectiveness of multi-visit screening

Many of the emerging alternative technologies and methods use an approach called multi-visit LDAR (MVL). In MVL, rapid, typically mobile 'screening' technologies take advantage of the heavy-tailed leak-size distribution to quickly identify the most problematic facilities. Close-range follow-up inspectors are then tasked to these facilities. In order for alternative programs to work, they must meet two requirements. Regulators require that these programs achieve equivalent mitigation, while producers require that they be less expensive than using only close-range methods, which are already approved by regulators. Although emerging technologies commonly propose to use MVL, this approach has never been shown to simultaneously meet both mitigation and cost requirements. This knowledge gap is addressed in Chapter 5.

1.4 Thesis overview

The overall objective of my thesis is to evaluate current and emerging methane-sensing technologies for emissions monitoring in upstream oil and gas. The knowledge gaps identified in the previous section are addressed in four publishable peer-reviewed articles, each a chapter of this thesis. This section contains a brief overview of each chapter.

1.4.1 Chapter 2

Chapter 2 is a literature review entitled "A review of close-range and screening technologies for mitigating fugitive methane emissions in upstream oil and gas." A version of this chapter was published in *Environmental Research Letters* in 2019 (Fox et al., 2019a). Chapter 2 provides a detailed review of what technologies and methods exist for monitoring methane emissions from oil and gas operations. We discuss the strengths and limitations of six technology classes: handheld instruments, unmanned aerial vehicles (UAVs), mobile ground labs (MGLs), piloted aircraft, satellites, and fixed sensors. We find that fixed sensors, UAVs, MGLs, and aircraft have potential to be deployed effectively in LDAR, but that satellites are not yet ready for monitoring of point-source emissions. Given that each technology class differs greatly in the conditions under which it excels, we introduce the notion of a comprehensive monitoring program, in which different technologies are deployed at different spatial and temporal scales.

1.4.2 Chapter 3

Chapter 3 presents the world's first proposed framework for demonstrating equivalence among alternative and regulatory LDAR programs. A version of this chapter was published in *Elementa: Science of the Anthropocene* in 2019 (Fox et al., 2019b). Chapter 3 is the result of a set of three interactive workshops in Calgary and Fort Collins with stakeholders from academia, industry, government (federal, provincial, and state in Canada and the US), consulting, technology development, and the non-profit sector. Together, these stakeholders iteratively and collaboratively developed a five-stage procedure for demonstrating equivalence: (1) Method identification, in which new methods are assembled and defined; (2) Controlled testing, in which the performance of new methods is evaluated; (3) Simulation modeling, in which computer models are used to predict alternative LDAR program performance; (4) Field trials, in which operational efficacy of new programs is established, and (5) Full approval by the regulator. Challenges to implementation are discussed, including controlled testing, modeling, scale and source disambiguation, human factors, and logistics.

1.4.3 Chapter 4

Chapter 4 introduces a novel simulation framework called the Leak Detection and Repair Simulator (LDAR-Sim) that enables comparison of simulated LDAR programs. A version of this chapter is in peer review (Fox et al., in review). LDAR-Sim is an open-source numerical model that estimates emissions from a variety of LDAR programs that consist of different technologies, methods, or regulations. Agents are deployed and sequentially visit facilities while constrained to real space and time. The modeling framework is modular and highly adaptable, enabling precise program definition with different facility types, simultaneous deployment of multiple different technologies, representative historical environmental conditions, scale differentiation using either close-range inspection or facility-scale screening, and different emissions source categories. We present several case studies of different LDAR programs in Alberta Canada to demonstrate the importance of weather, labour constraints, and modeling assumptions. We also show that screening technologies can underperform in the presence of legal methane venting.

1.4.4 Chapter 5

In Chapter 5 I use LDAR-Sim to explore whether MVL can achieve equivalence while being more cost-effective than regulatory LDAR. A range of 'equivalence scenarios' are developed, which represent a screening survey frequency combined with a follow-up requirement. We find that as survey frequency increases, the number of follow-up surveys required to achieve equivalence decreases. An important step in MVL is identified, called triaging, which is a decision-making procedure that determines which of the screened facilities should receive follow-up inspection with close-range instruments. A common form of triaging involves ranking facilities by their emission rate, but we find that the ranking order can be incorrect if screening methods have high quantification error or if vented emission sources are present. Incorrect ranking can in turn lead to sending follow-up inspectors to the wrong facilities, which increases emissions and requires additional surveys and greater cost to maintain equivalence. Our analysis finds that under most realistic scenarios, MVL is unlikely to be a cost-effective replacement for regulatory LDAR at this time. Further innovation in this space or regulatory changes may improve MVL cost-effectiveness and appeal.

Chapter 2

A review of close-range and screening technologies for mitigating fugitive methane emissions in upstream oil and gas

2.1 Abstract

Fugitive methane emissions from the oil and gas industry are targeted using leak detection and repair (LDAR) programs. Until recently, only a limited number of measurement standards have been permitted by most regulators, with emphasis on close-range methods (e.g. Method-21, optical gas imaging). Although close-range methods are essential for source identification, they can be labour-intensive. To improve LDAR efficiency, there has been a policy shift in Canada and the United States towards incorporating alternative technologies. However, the suitability of these technologies for LDAR remains unclear. In this paper, we systematically review and compare six technology classes for use in LDAR: handheld instruments, fixed sensors, mobile ground labs (MGLs), unmanned aerial vehicles (UAVs), aircraft, and satellites. These technologies encompass broad spatial and temporal scales of measurement. Minimum detection limits for technology classes range from < 1 g/h for Method 21 instruments to 7.1 x 10^6 g/h for the GOSAT satellite, and uncertainties are poorly constrained. To leverage the diverse capabilities these technologies, we introduce a hybrid screening-confirmation approach to LDAR called a comprehensive monitoring program (CMP). Here, a screening technology is used to rapidly tag high-emitting sites to direct close-range source identification. Currently, fixed sensors, MGLs, UAVs, and aircraft could be used as screening technologies, but their performances must be evaluated under a range of environmental and operational conditions to better constrain detection effectiveness. Methane-sensing satellites are improving rapidly and may soon be ready for facility-scale screening. We conclude with a speculative discussion of the future of LDAR, touching on integration, analytics, incentivization, and regulatory pathways.

2.2 Introduction

Methane, a potent greenhouse gas (GHG), has become a major focus of GHG reduction initiatives. Of the many sources of natural and anthropogenic methane, emissions from the oil and gas (O&G) industry have received special attention (Moore et al., 2014). In Canada, commitments have been made to cut methane emissions from the O&G sector by 40 – 45% below 2012 levels by 2025 (ECCC, 2018). In addition to meeting GHG targets, reducing methane emissions from O&G can save money (ICF International, 2015, 2014) and improve local air quality (Roy et al., 2014).

Natural gas (NG) consists primarily of methane and is invisible and odourless in most upstream settings. Within regulatory and operational contexts, releases of NG to the atmosphere are often classified as either vented or fugitive emissions. Vented emissions are intentional releases of hydrocarbons, typically in a controlled manner, resulting from normal process conditions. In contrast, fugitive emissions (also called 'leaks') are unintentional releases of hydrocarbons from sources that should not be emitting (e.g. broken valves, flanges, etc.) In general, NG emissions exhibit high spatial and temporal variability (Heimburger et al., 2017; Lavoie et al., 2017; Robertson et al., 2017) and are difficult to predict (Brandt and Pétron, 2015; Kemp et al., 2016). Across North America, sources exhibit highly-skewed leak-size distributions, with a small number of 'super-emitters' accounting for a disproportionate share of total emissions (Brandt et al., 2014; Zavala-Araiza et al., 2017, 2015).

To mitigate fugitive emissions, operators commonly implement leak detection and repair (LDAR) programs. Method 21 (Determination of Volatile Organic Compound Leaks), introduced in 1983 by the US Environmental Protection Agency (EPA), was the first regulatory framework for conducting LDAR. In 2008, the EPA expanded the scope of possible compliance procedures by introducing the Alternative Work Practice (AWP). AWP replaces concentration detectors with optical gas imaging (OGI) cameras – handheld instruments that generate infrared images of methane plumes. Today, OGIs are preferred over Method 21 by regulators and operators in Canada and the US, and most LDAR programs rely exclusively on handheld methods, which can be labour-intensive. Both Method 21 and AWP require facility access, and LDAR must be applied to millions of components distributed over extensive spatial scales. For example, Figure 2.1 illustrates the variable density of O&G infrastructure across the Canadian provinces of Alberta and British Columbia, with densities often below two wells per square kilometre. At each facility, LDAR technicians must manually survey hundreds or thousands of individual components.

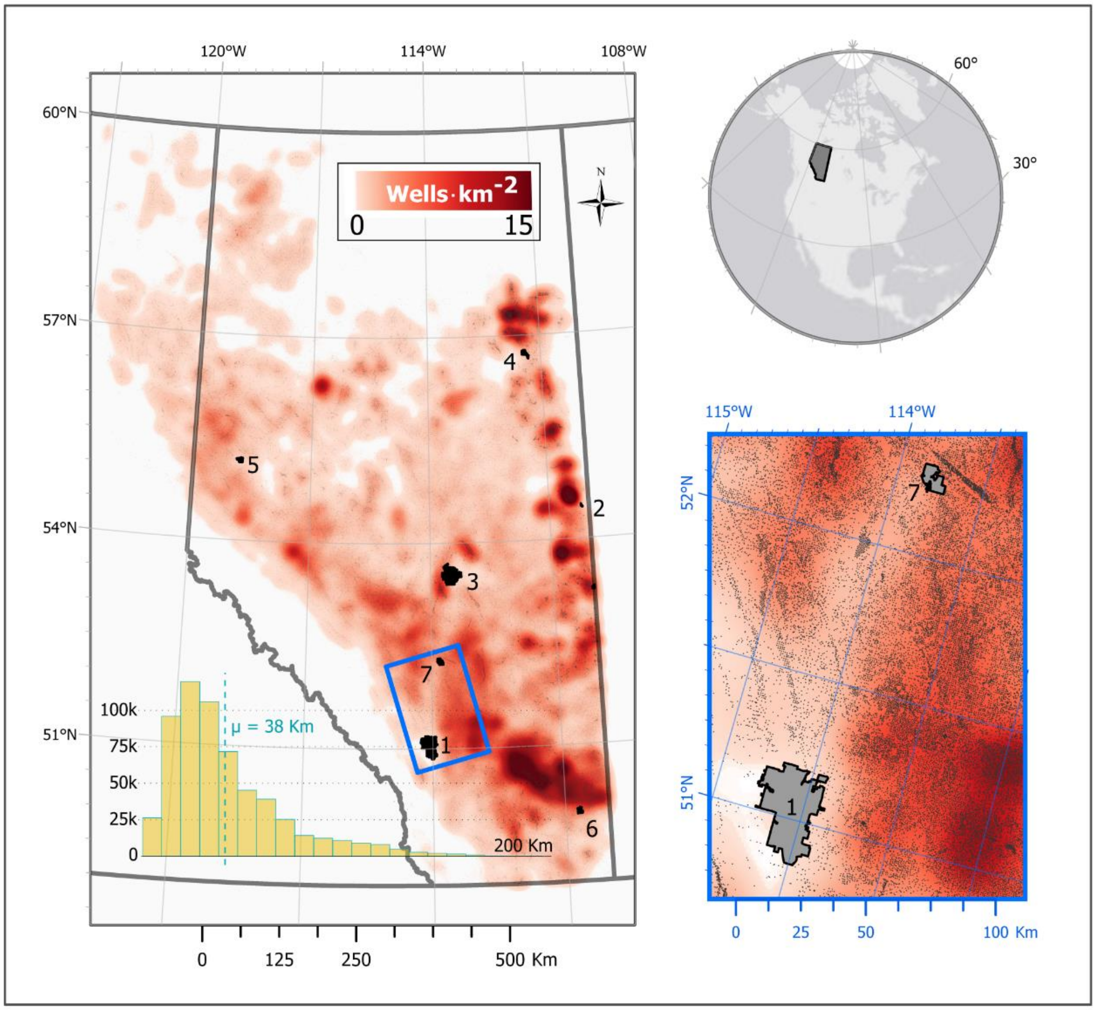

Figure 2.1 Kernel density map showing spatial variability of well density in Alberta and Northeastern British Columbia, Canada. Important oil and gas municipalities include 1) Calgary; 2) Cold Lake; 3) Edmonton; 4) Fort McMurray; 5) Grande Prairie; 6) Medicine Hat; and 7) Red Deer. Inset histogram represents the distance from each well to the nearest of 1010 population centres (μ = 38 km, range = 0-290 km). Well data was acquired from the Alberta Energy Regulator in June 2017.

Given the spatial extent and variable density of O&G infrastructure, deploying rapid screening technologies may improve LDAR efficiency. Screening technologies are systems that can quickly flag abnormally emitting facilities for directed follow-up with close-range methods. In recent decades, there have been major developments in sensors, platforms, and analytics that could serve a screening role, including mobile ground labs (MGLs), fixed sensors, aircraft, unmanned aerial vehicles (UAVs), and satellites. Typically, screening technologies can neither identify leaks at the component-level, nor distinguish vented from fugitive emissions. To diagnose and repair leaks, screening methods must be paired with close-range methods such as Method 21 and AWP, which can precisely pinpoint leaking components. We refer to a program that combines screening and close-range methods as a comprehensive monitoring program (CMP).

Regulators are moving towards policy frameworks that enable CMPs. In April 2018, Environment and Climate Change Canada finalized new regulations allowing operators to choose between prescriptive "Regulatory" LDAR and "Alternative" LDAR (ECCC, 2018). In Regulatory LDAR, operators must use Method 21 or OGIs to monitor for fugitive emissions at least three times per year at select facilities. In Alternative LDAR, operators are invited to develop LDAR programs incorporating new methods and technologies. To be approved, Alternative programs must demonstrate emissions reductions equivalent to Regulatory LDAR. In Alberta, regulations released in December 2018 mandate screening for fugitive emissions using one of several methods, including UAVs and truck-mounted sensors (i.e. MGLs). In the US, measures for the approval of alternative LDAR have been implemented in 40 CFR Part 60 Subpart OOOOa (EPA, 2016). Similar opportunities are outlined in Section XII.L.8 of Colorado's Regulation 7 (CDPHE, 2018).

Screening technologies are under various stages of development and commercialization and there are gaps in the information available to guide deployment and regulation. As a growing number of technology developers promise solutions, regulators and operators must determine how (and whether) to integrate these technologies into current LDAR programs. In this paper, we systematically evaluate and compare leak detection technologies, and discuss ways in which LDAR programs might integrate these technologies to mitigate emissions. This article is presented in 6 sections. In Section 2.3 we establish the methodological framework for the review and discuss measurement techniques and quantification. Section 2.4 is an overview of methane measurement principles and the platforms used for monitoring methane in O&G. In Section 2.5 we review current and emerging methane-sensing technologies in the context of LDAR. Section 2.6 is a speculative discussion of the future of LDAR, touching on technology evaluation, multiscale integration, and regulatory models. Section 2.7 is a brief conclusion.

2.3 Review methods and framework

Our review is focused on six broad technology classes: handheld instruments, fixed sensors, mobile ground labs (MGLs), unmanned aerial vehicles (UAVs), aircraft, and satellites. We searched Google Scholar and the University of Calgary library database using the keywords: leak detection and repair, fugitive methane emissions, and methane sensing oil and gas. Relevant articles were then used to identify further sources, by consulting both ‘works cited’ and ‘cited by’ lists. This process was repeated until no new sources could be identified. Non-peer-reviewed sources were sometimes used, such as government reports or independent research publications. Sources published after 25 October 2018 may not be included.

Although this review is focused primarily on LDAR and screening for anomalous emissions, it is useful to consider whether candidate technologies are suitable for different monitoring programs. To frame this review, we consider four distinct motivations for measuring NG emissions from upstream O&G:

M1) Develop and refine emissions factors to improve inventories,

M2) Estimate top-down emissions from a region with multiple sources,

M3) Conventional, close-range LDAR using handheld instruments, and

M4) Rapid screening for anomalous emissions.

These motivations stem from two fundamental goals:

Goal 1: Understand emissions (M1 and M2), and

Goal 2: Mitigate emissions (M3 and M4).

For each goal, equipment can be targeted at a granular scale (M1 and M3), or at an aggregate scale (M2 and M4). Different technologies and methods are required for each motivation, and different data products can be expected. For example, developing emissions factors requires accurate quantification, often at the component-level. In contrast, estimating top-down emissions requires mobile (often airborne) platforms capable of resolving small concentration enhancements dozens of kilometres downwind of a source region. Close-range methods often favour real-time imaging and generally do not require quantification. Screening should be less expensive than close-range methods, but deployment strategies can differ. First, screening methods can inform directed application of follow-up surveys. For example, a close-range survey of a facility could be skipped if there are no anomalous emissions identified through screening. Second, among detected emissions sources, screening methods can help triage follow-up and repair based on a size-ordered list of flagged facilities, reducing aggregate emissions as the largest leaks are repaired first. Third, screening methods can focus on super-emitter targeting. Given skewed leak-size distributions, early identification of super-emitters could mitigate a majority of emissions. In super-emitter targeting, screening methods should have low per-site cost, high spatial coverage, and rapid return intervals. If a field contains few super-emitters (i.e., a less-skewed leak size distribution), implementing super-emitter targeting may be less effective.

We evaluate technologies across three product levels: 1) detection, 2) localization and/or attribution, and 3) quantification. For close-range methods, detection and localization are often accomplished simultaneously, and quantification is generally less important. For screening, quantification is often necessary to determine whether a follow-up survey should be conducted using close-range methods. For technologies with high detection limits, quantification could be less important, as each detection event could trigger a follow-up survey. If multiple detection events occur during screening, relative quantification can enable triaging. Quantification may also permit the separation of vented from fugitive emissions, but only where vented emissions are precisely known.

For LDAR in general, quantification may be important depending on the goals of the program. If mitigating fugitive emissions is the primary goal, quantification may be less important than detection, as quantification generally takes more time and money that could instead be invested in more frequent detection-only surveys. However, if the goal is to conduct LDAR while developing an improved scientific understanding of the root causes of emissions sources, to reduce uncertainty in inventories, or to track progress in emissions reduction initiatives, quantification and data management become increasingly important. Quantification may further help by improving accountability and trust among industry, government, and the public.

2.4 Technology overview

2.4.1 Measuring methane

Today, most methane concentration measurements are made with optical instruments, using either laser spectroscopy or imaging spectrometry. Laser spectroscopy determines the concentration of target molecules by measuring characteristic absorption of a mid- or near-infrared laser along a path length of metres to kilometres. The laser path may be “open,” where it goes through the immediate atmosphere, or “closed,” using a mirrored cavity into which gas is pumped. Unlike laser-based instruments, imaging spectrometers measure spectral densities using pixel-based sensor elements. For methane, abundance can be inferred using specific infrared absorption bands. Imaging spectrometers generate a multi-pixel field of view measurement that captures column-integrated concentrations. Other sensor classes exist, such as ionization devices and differential absorption light detection and ranging (DIAL). An understanding of sensor differences is useful for comparing technologies, but a detailed description of gas-sensing principles is beyond the scope of this article.

Once a screening technology acquires concentration data, atmospheric dispersion models can help determine source location and emission rate. These vary in complexity from simple Gaussian dispersion models to complex particle-tracing Lagrangian models that account for turbulence. Results using both simple and complex dispersion models have shown good equivalency in methane-specific studies (Brantley et al., 2014; Foster-Wittig, 2015; Lan et al., 2015). Dispersion modelling has been used for O&G monitoring of pollutants for decades, and most regulators publish guidance documents mandating what models to apply and how. However, most established techniques were designed and validated for stationary sensing, and it remains unclear how transferrable they are to mobile platforms. Recent studies comparing different quantification methods suggest that more work is needed to constrain quantification uncertainties (Bell et al., 2017; Caulton et al., 2018).

2.4.2 Technology classes

The technologies reviewed in the following section encompass broad spatial and temporal scales of measurement (Figure 2.2). In Figure 2.2, the spatial scale refers to the order of magnitude length-scale typically covered during a single measurement campaign, while the temporal scale refers to measurement times over which a single survey is completed once the technology is deployed. For example, satellite-based monitoring systems can measure across the planet, quasi-continuously, for many years (large spatial and temporal scales). Conversely, close-range instruments such as OGIs may only cover a few kilometres, over the course of a few hours to a few days (small spatial and temporal scales). Fixed sensors tend to monitor at large temporal and small spatial scales. As technologies evolve, the size and position of the ellipses in Figure 2.2 may change. Currently, the dotted horizontal line at y = ‘day’ divides semi-automated (above) and labour-based systems (below). As a hypothetical example, if MGLs become driverless their range could expand into broader spatial and temporal scales. Similarly, the UAV niche couple expand if battery limitations are overcome (i.e. with solar power) and airspace regulations relax to permit UAV operations beyond visual line of sight. In the future, a satellite cluster with higher spatial resolution and revisit time of less than a day could move satellites towards continuous monitoring.

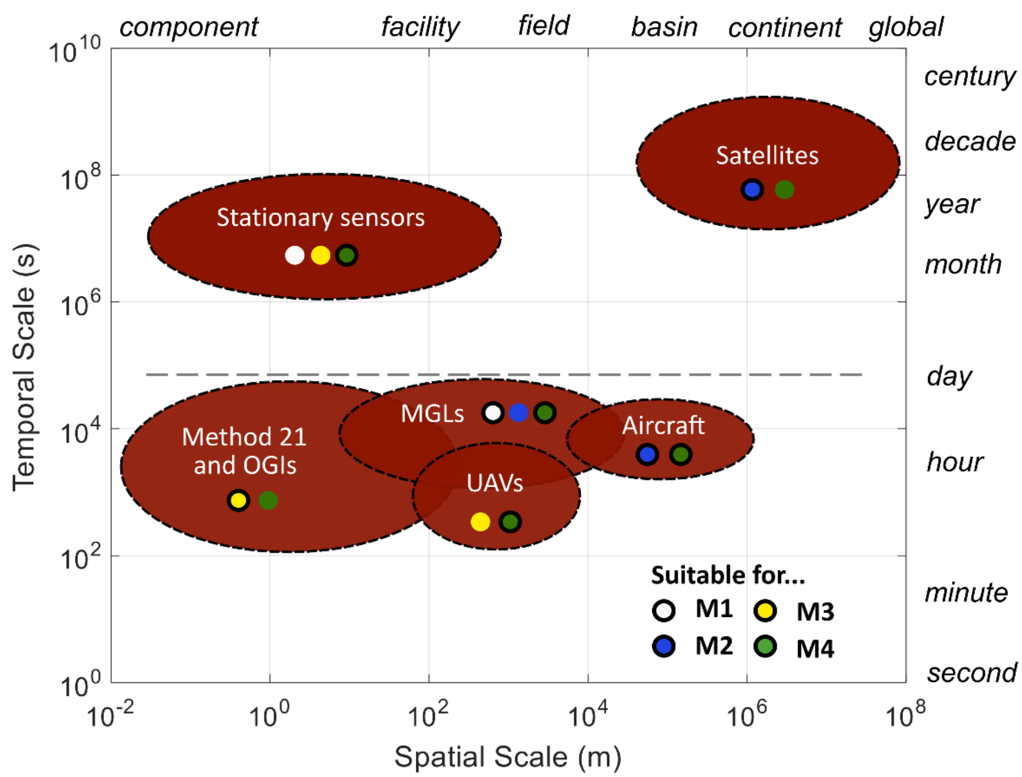

Figure 2.2 Technology classes categorized based on the spatial and temporal extent of coverage. Coloured dots represent suitability for measurement motivations 1-4. Dots without black borders either show promise or may be useful in a limited capacity.

A summary of the technology metrics used to structure this review is presented in Table 2.1. Certain themes were omitted from the table if limited information was available or if the metrics were only applicable to specific technologies. These themes are discussed in the text and include operational conditions (e.g. susceptibility to adverse meteorology), technology readiness levels for the four monitoring motivations introduced in Section 2.3, and future potential. We note that the quantitative comparison of technology classes presented in Table 2.1 must be interpreted with caution for three reasons. First, most sensors and platforms have not been sufficiently evaluated under a range of environmental and operational conditions, and only a handful of studies have reported detection limits and uncertainties for each technology class. Published measurement ranges may not be representative, and context-dependent detection probability distributions should be developed before different technologies can be reliably compared. Second, sensors, platforms, and analytics are all under active development, with evolving technical capabilities and modes of deployment. Third, incompatible sensing principles (e.g. point concentration vs. integrated path-length measurements) and modelling approaches make quantitative comparisons difficult. Developing technology equivalence protocols would enable a systematic comparison of fundamentally different technologies; such an effort is beyond the scope of this article.

An important metric that was not quantitatively included in this study is cost. The costs of an LDAR program include equipment purchase or rental, labour, repair, and additional considerations such as insurance, training of personnel, and equipment maintenance. Program costs can be compared on a per facility basis, or cost per unit of mitigated emissions. Unfortunately, too few studies have reported deployment costs to enable a quantitative comparison of technology classes. Although estimates have been developed for OGIs (e.g., ICF International 2014; 2015; Saunier, Haugland, and Pederstad 2014; various regulatory estimates) and aircraft (Schwietzke et al., 2018), more data are needed for these and other technologies, and a comprehensive economic analysis is beyond the scope of this review.

2.5 Technology review

2.5.1 Handheld instruments

The most common sensors used for Method 21 are flame and photoionization detectors, although catalytic oxidation sensors, and infrared absorption-based sensors, are also used (Envirotech Engineering, 2007; Szulczynski and Gebicki, 2017). Although detection limits tend to be low, Method 21 instruments are labour-intensive as the sensing probe should be in immediate proximity to the leaking component for detection (Table 2.1). Method 21 is still favoured by some operators but use is declining as OGIs are more convenient. Furthermore, flame ionization detectors are not intrinsically safe for use in O&G.

In recent years, OGIs have become the standard for LDAR because they generate easily communicable and intuitive results for reporting purposes, and are more efficient than Method 21 (Table 2.1), as they survey components remotely (Benson et al., 2006; Reese et al., 2006). Thus, OGIs are capable of limited screening, which is restricted by imaging distance to small spatial scales (Ravikumar et al., 2018). Despite their simplicity and widespread use, OGIs have some limitations. First, their operation is labour-intensive, as technicians should be within a few meters of potential sources. Recent work suggests that camera-to-source distance is the most important factor in predicting OGI detection effectiveness (Ravikumar et al., 2017), and other work shows that imaging distances in excess of ~10 m suffer from significantly reduced performance (Ravikumar and Brandt, 2017). Surveying thousands of wells, from within 10 m, several times per year, could result in considerable operational cost (ICF International, 2015, 2014).

Table 2.1 Comparison of leak detection technologies and methods.

	Method 21	OGIs	Fixed Sensors	MGLs - Stationary[a]	MGLs - Tracer[a]	MGLs - Mobile	UAVs	Aircraft - Facility-scale	Satellites - Facility-scale
Limit of detection (g/hr)[b]	< 1[c]	20[d]	96[e]	9 - 36	700 - 1.2 X 10^4	6 - 2124	39.6[f]	2000 - 4.6 X 10^4	2.5 X 10^5 - 68 X 10^6 [g]
Flux estimation uncertainty (%)[h]	-	3 - 15[i]	31	25 - 60	20 - 50	50 - 350	25 - 55	1 - 24	j
Horizontal distance from source (m)	0	3 - 6	0 - 1000	10 - 200	500 - 3000	5 - 500	0 - 194	0 - 1000	-
Vertical distance from source (m)	-	-	-	-	-	-	6.5 - 122	60 - 1000	5.12 X 10^5 - 8.24 X 10^5
Time per well pad (minutes)[k]	240 - 960[l]	120 - 480	-	15 - 20	60 - 300	0.5 - 5	5 - 15	5 - 30	< 0.01
Sensor in plume?	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No
Component-level confirmation?	Yes	Yes	No	No	No	No	No	No	No
Regulatory acceptance for LDAR?[m]	Yes	Yes	No	No	No	No	No	No	No
Readiness level for LDAR	High	High	Moderate	-	-	Moderate	Moderate	Moderate	Low
Number of operational commercial systems for use in LDAR[n]	20+	5+	20+	-	-	10+	20+	10+	1
References and important publications	1 - 3	3 - 7	8 - 9	10 - 12	13 - 15	16 - 20	21 - 23	24 - 28	29 - 30

[a] Deemed unsuitable for screening, see section 4.3. [b] These are examples of specific studies under specific conditions; more research is needed. [c] Instrument-dependent. [d] 90% probability of detection at 3 m. [e] At 1 km. [f] For measurements taken within 10 m. [g] Theoretical detection limits. [h] This row is largely adapted from 17. [i] OGI quantification is complex and uncertainties are likely much higher. [j] There are no published uncertainty estimates for facility-scale emissions. [k] Measurement time only. [l] Assuming Method 21 is 50% as efficient as OGI. [m] In Canada/US. [n] Examples from a quick web search; more may exist or be in development. Companies are not listed to avoid endorsement.

1 Ellis and Lackaye 1989 **2** Yen and Horng 2009 **3** Ravikumar et al. 2018 **4** ICF International 2014 **5** Gålfalk et al. 2016 **6** Ravikumar and Brandt 2017 **7** Ravikumar, Wang, and Brandt 2017 **8** Coburn et al. 2017 **9** Patel 2017 **10** Brantley et al. 2014 **11** Lan et al. 2015 **12** Robertson et al. 2017 **13** Mitchell et al. 2015 **14** Roscioli et al. 2015 **15** Yacovitch et al. 2017 **16** Yacovitch et al. 2015 **17** Caulton et al. 2018 **18** Atherton et al. 2017 **19** von Fischer et al. 2017 **20** Weller et al. 2018 **21** Nathan et al. 2015 **22** Barchyn et al. 2017 **23** Golston et al. 2018 **24** Frankenberg et al. 2016 **25** Conley et al. 2017 **26** Smith et al. 2017 **27** Englander et al. 2018 **28** Schwietzke et al. 2018 **29** Frankenberg et al. 2005 **30** Jacob et al. 2016

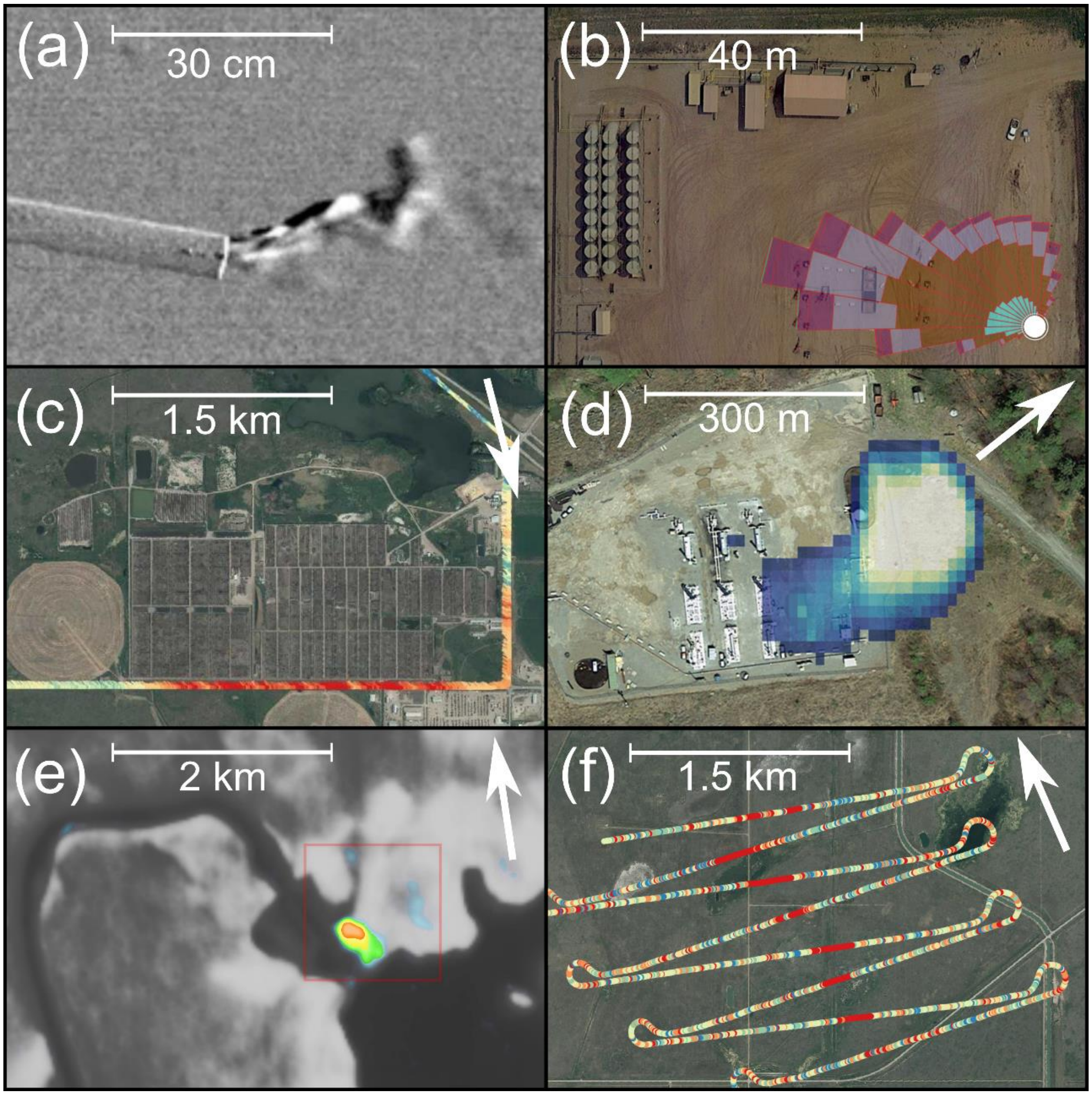

Figure 2.3 Examples of data output scale and form: a) FLIR GF320 image of an emitting pipe; b) Methane flux wind rose of a fixed sensor (provided by Quanta3); c) MGL concentration map near a large cattle feedlot; d) Piloted aircraft retrieval using imaging spectroscopy (provided by Kairos Aerospace); e) Column-averaged methane enhancements from a flooded dam (provided by GHGSat); and f) fixed-wing UAV concentration map of a controlled release. White arrow indicates wind direction.

OGI performance is also affected by adverse environmental conditions such as high wind speeds (Footer, 2015), low ambient air temperatures, and low background emissivity contrast (Ravikumar and Brandt, 2017). While environmental constraints might be less of a concern in warmer climates, the majority of O&G infrastructure at high latitudes (e.g. Canada, Norway, Siberia, etc.) experiences months of below freezing temperatures, often accompanied by high wind speeds. Operator expertise also plays a role in detection effectiveness (Footer, 2015; von Fischer et al., 2016), meaning that appropriate training, compliance auditing, and incentivization structures are needed. Finally, most current OGIs only present a qualitative (visual) flux estimate. Recent software products claim to estimate flux rate from OGI videos of leaks, though their effectiveness is yet to be established in the peer-reviewed literature (Providence Photonics, 2018; Rebellion Photonics, 2018).

Method 21 instruments and OGIs are not currently able to contribute to motivations M1 and M2, can provide limited support for M4, and are vital for M3. Despite their limitations, OGIs are the most widely-used technology for LDAR, and are unlikely to be replaced by screening technologies because component-level localization is critical for repair and reporting. OGIs are well-suited for detecting large plumes under favourable imaging conditions. Unlike screening technologies, they also provide exact, component-level localization. Current North American policies support the use of OGIs, so it is reasonable to expect that competition in a growing market will drive sensors to be 1) less expensive, 2) more reliable under unfavourable measurement conditions, and 3) better able to quantify flux. New technologies, like hyperspectral cameras (Gålfalk et al., 2016) and laser-based handheld spectrometers (Wainner et al., 2017), may offer improved detection and quantification at lower cost.

2.5.2 Fixed sensors

Fixed sensors are deployed in high-risk areas and provide continuous readings of methane concentration, triggering an alarm should concentrations exceed a predefined level. To date, optical methods are most common, including laser-based line-integration sensors, fixed concentration detectors, and camera installations. However, potentially lower-cost solid-state sensors are under active development (ARPA-E, 2018; Patel, 2017). Similar to OGIs, camera installations are larger, more accurate, and permanently mounted (Gerhart et al., 2013). Line integration sensors use near- or mid-infrared laser spectroscopy to measure methane (Goldsmith et al., 2012; Hashmonay and Yost, 1999; Ro et al., 2011). A laser travels from a sensor to a retro-reflector, and then returns to a photo-diode. These sensors are well suited for permanent installations, but most have path lengths limited to 100s of meters (Coburn et al., 2017). For monitoring two-dimensional areas, some line integration sensors use beam splitters to measure across different paths. There have also been scientific and commercial developments of line-integration sensors with longer path-lengths (> 1 km), capable of localizing and quantifying methane emissions over large areas (Alden et al., 2017; Coburn et al., 2017). Stationary systems have been successfully developed and evaluated for CO2 emissions monitoring (Hirst et al., 2017).

Fixed sensors could potentially contribute to M1 and M2 and already contribute to M4. Continuous monitoring and potential for automation make fixed sensors appealing, especially in dense infrastructure. As a screening tool, a distributed sensor network could identify fugitive emissions nearly instantaneously, preventing extended emissions events that remain undetected between mobile screening and conventional LDAR visits. As the only non-mobile technology class, fixed sensors might be best suited for facilities with high component density (e.g. gas plants, compressor stations, multi-well pads). For sparsely distributed upstream O&G infrastructure, these sensors must either be affordably mass-produced, or able to monitor large areas. Some progress has been made in developing low-cost sensors (e.g. Patel 2017). The ARPA-E MONITOR program funded several projects that promise cost-effective solutions for detection, localization, and quantification. These include distributed sensor networks costing as low as $300/sensor, printed carbon nanotube sensors with 1 ppmv sensitivity, and affordable mid-infrared integration networks, among others (ARPA-E, 2018). However, many of these technologies are still in development and are undergoing field trials; it is not known whether they will perform according to expectations. Few fixed sensors have been independently evaluated for hydrocarbon leak detection in upstream O&G, and technical capabilities and limitations of this technology class are poorly constrained (Table 2.1). Nevertheless, the appeal of continuous, automated monitoring has the potential to catalyse progress in this area.

2.5.3 Mobile ground labs

MGLs are versatile platforms for conducting local- to regional-scale surveys of methane emissions. In their simplest form, MGLs consist of a vehicle equipped with a global positioning system and a methane sensor. This setup enables a survey approach called concentration mapping (Figure 2.3c), which generates a map of methane concentrations along the vehicle path (Atherton et al., 2017; von Fischer et al., 2017). Simultaneous measurement of methane and a second thermogenic species (e.g. ethane) can improve plume characterization (Atherton et al., 2017; Yacovitch et al., 2015). Biogenic and thermogenic sources can also be distinguished using isotopic analysis (Townsend-Small et al., 2012; Zazzeri et al., 2015). As a screening tool, MGLs may be used in close proximity (e.g. on the well pad) or on nearby roads, up to several kilometres downwind, but with detection limits increasing with distance (von Fischer et al., 2017). Onsite screening can detect smaller sources but may miss others that are elevated, indoors, or otherwise inaccessible. Onsite and road-based screening methods have not been quantitatively compared. To date, mobile quantification studies have all relied on offsite measurements. As detected plumes generally originate upwind from roads, dispersion models must be used for localization and quantification estimates (Yacovitch et al., 2015). Numerous models have been developed to remotely estimate mass flux from an MGL (Section 2.5.3), including EPA Other Test Method (OTM) 33A (Brantley et al., 2014; Thoma and Squier, 2014), variations on the Gaussian plume model (Lan et al., 2015; Yacovitch et al., 2015), the mobile flux plane technique (Rella et al., 2015), or statistical techniques (von Fischer et al., 2017). If site access is available, and time permits, tracer techniques may also be used (Mitchell et al., 2015; Roscioli et al., 2015; Yacovitch et al., 2017), but these are unlikely to be useful for screening due to per-site time requirements (Table 2.1).

Passive MGL measurements are made by mounting instruments on vehicles performing unrelated tasks (Albertson et al., 2016; Christen, 2014). For example, Google Street View vehicles were recently used to measure natural gas leaks in distribution systems in the US cities of Boston, Indianapolis, New York, Syracuse and Burlington (von Fischer et al., 2017). The vehicles were tasked with taking photographs of street scenes, but the addition of a methane sensor allowed them to also collect methane concentration and identify locations of thousands of urban leaks. This approach is advantageous due to low implementation costs, although it is more difficult to collect data in certain locations, as routes are slaved to the primary task. Passive sensing and road-dependency make MGLs well-suited for urban settings. These capabilities have already proven to be useful for identifying emissions sources in municipalities with older natural gas distribution infrastructure, such as Washington D.C. (Jackson et al., 2014) and Boston (Phillips et al., 2013). Whether for rural or urban applications, most modern MGLs are equipped with a user interface that displays real time concentration data, enabling drivers to search for, locate, and repair simultaneously.

In recent years, MGLs have received considerable interest and have been deployed in various ways (Table 2.1) and could contribute to M1 (e.g. using the tracer flux technique), M2 (e.g. with passive monitoring), and M4 (using mobile methods). Some approaches require that MGLs be stationed within the plume of interest for upwards of 15 minutes (e.g. OTM33A), but newer, less-precise methods use in-motion quantification (Table 2.1; Albertson et al. 2016; Rella et al. 2015; Yacovitch et al. 2015). Although detection limits for mobile MGLs are higher, have greater uncertainties, and may underestimate small sources, they may be more appropriate for screening as the time spent at each facility is < 5 minutes, compared to a minimum of 15 minutes for OTM33A and 60 minutes for tracer methods (Table 2.1; Weller et al. 2018). Recent work, in which over 1600 well pads were surveyed across nearly 8000 km of roads, has demonstrated the potential of MGLs to screen large areas (Atherton et al., 2017). Approaches that do not require site access spend less time at each site, require minimal coordination with facility operators, and provide enforcement agencies and independent researchers the benefit of a blind sample. Despite these advantages, MGLs are unlikely to meet all the needs of an LDAR program. First, MGLs are limited by road access and meteorological conditions, especially wind direction. In the absence of sufficient wind, or if wind is blowing in the wrong direction, fugitive plumes may never reach a road. Second, screening-grade quantification remains difficult. Plume lofting due to atmospheric instability is a challenge, although some attempts have been made to better characterize vertical concentration gradients (Rella et al., 2015). As with other screening methods, differentiating between routine venting and fugitive emissions remains unsolved, and criteria must be established for whether a measured enhancement warrants follow-up with close-range methods.

2.5.4 Piloted aircraft

In recent years, there has been a growing interest in using piloted aircraft for surveying site-level emissions. While helicopters may be used (Babilotte et al., 2010; Englander et al., 2018; Lavoie et al., 2017; Lyon et al., 2016), small airplanes can cover longer distances and fly numerous repeat sampling transects in three-dimensional space (Peischl et al., 2018; Schwietzke et al., 2018). Aircraft can be equipped with sensors that sample air at precisely known times and locations (e.g. Conley et al. 2016) or with imaging spectrometers that generate column-averaged methane concentrations (Figure 2.3d; e.g. Buckland, Burnham, and Augustin 1997; Frankenberg et al. 2016). Aircraft that process air samples use instruments similar or identical to those used in MGLs. However, MGLs may only collect data on a two-dimensional surface, while aircraft can resolve methane plumes in three dimensions, and do not require roads.

Aircraft can contribute to M2 and M4. Historically, M2 has been more common, typically in the form of the mass balance approach (Johnson et al., 2017; Karion et al., 2015, 2013), which measures methane concentration around a targeted source area and attributes the difference between upwind and downwind measurements to mass flux contributions from the region of interest (Butler et al., 2004). However, the mass balance approach provides little value for LDAR and mitigation, as individual facilities are not resolved. Cylindrical mass balance approaches are an evolution of traditional methods that may target facility-scale sources (Conley et al., 2017; Smith et al., 2017; Tadić et al., 2017). Multiple-line airborne surveying, which can also be used to map and quantify emissions, is an alternative to the mass balance approach specifically designed to identify high-emitting facilities (Hirst et al., 2013; Terry et al., 2017). Imaging spectrometers can also be used for screening purposes (Thorpe et al., 2017). Some have used hyperspectral remote sensing to characterize individual methane plumes (Buckland et al., 2017; Frankenberg et al., 2016). A recent study simultaneously conducted regional and facility-scale surveys using aircraft mass balance (Smith et al., 2017), and the use of path-integrated DIAL has shown promise (Bartholomew et al., 2017). Compared to remote sensing, air sampling methods currently have lower detection limits by approximately one order of magnitude and are more cost-effective than imaging approaches for mitigating emissions (Schwietzke et al., 2018). Although the hyperspectral AVIRIS-NG has detection limits comparable to the cylindrical mass balance method, this research-grade instrument may currently be too expensive to scale commercially for LDAR (Schwietzke et al., 2018; Thorpe et al., 2017).

For screening, the main strength of aircraft relative to ground methods and UAVs is survey speed. Although satellite coverage is better, detection limits for aircraft are currently 1-3 orders of magnitude lower (Table 2.1). However, piloted aircraft have several features that may limit their use for fugitive emissions screening. In addition to the relatively high cost of acquisition, maintenance, and operation, aircraft are subject to regulations that dictate when and where surveys can occur. The most limiting of these regulations is the lowest safe altitude (LSA) for flying, which is generally a minimum of 500 feet above the highest regional structure. Therefore, despite having the ability to resolve methane plumes in three dimensions, issues may result if a significant proportion of the gas remains below the LSA. At high latitudes, the stable atmospheric conditions that lead to insufficient vertical mixing can prevail for months. In cases where some mixing occurs, efforts have been made to account for missing concentration data from below the LSA (Conley et al., 2017). It should be noted that this limitation only affects sampling approaches that pass through the plume, and that remote sensing techniques are not impacted, suggesting that the latter approach provides important advantages when the planetary boundary layer is shallow (Frankenberg et al., 2016). However, imaging spectrometers require adequate insolation, which can be limited by clouds and low radiative flux during high-latitude winters, depending on the region and time of year (Fig. 4). In general, aircraft have difficulty monitoring in winter conditions; radiation inversions prevent air sampling and scattering by snow confounds remote sensing based on infrared spectroscopy. In high-latitude production areas, aircraft surveys may not be possible for several consecutive months each year. Given the necessary distance between piloted aircraft and potential methane sources, minimum detection limits tend to be much higher than with ground-based methods (Table 2.1). As background methane concentrations are spatially dependent (Goetz et al., 2017; Verhulst et al., 2017), high aircraft velocities mean that considerable care must be taken to avoid unnecessary errors in the calculation of methane enhancements (Hirst et al., 2013). Accounting for the high temporal variability of O&G and non-target emissions is also challenging (Allen et al., 2017; Vaughn et al., 2018). Evidence suggests that even with numerous repeat flights it is difficult to characterize temporal variability in mass flux rates (Heimburger et al., 2017; Nathan et al., 2015; Schwietzke et al., 2017).

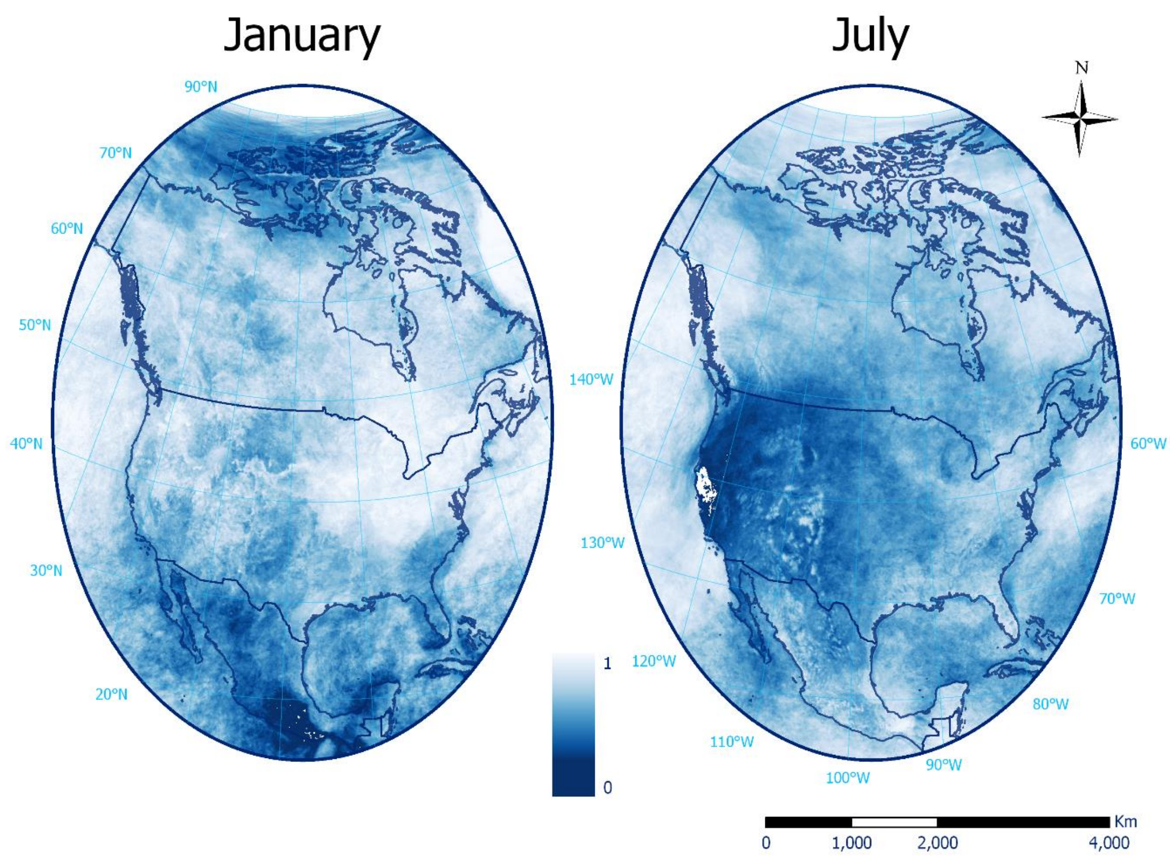

Figure 2.4 Monthly average NASA cloud fraction (Terra/MODIS, 0.1 degree) of North America for January and July 2017. Cloud cover can limit the application of aircraft- and satellite-based imaging spectrometers.

2.5.5 Unmanned aerial vehicles

UAVs have received considerable interest for their use in LDAR (Figure 2.3f). Like aircraft, UAVs can operate in three-dimensional space and are not restricted to roads, but have significantly lower detection limits (Table 2.1). Compared to aircraft, UAVs are also cheaper and more flexible to operate, are semi-automated, and can complement close-range methods by reaching dangerous or inaccessible places, such as the tops of tanks. One of the primary advantages of UAVs is that they are uniquely suited to characterizing methane plumes in three dimensions while flying near the source, improving confidence in attribution. To date, various attempts have been made to measure methane from a UAV (Barchyn et al., 2017; Brownlow et al., 2016; Cossel et al., 2017; Emran et al., 2017; Khan et al., 2012; Smith et al., 2017). The first comprehensive study investigating UAV methane-sensing potential focused on a single compressor station, which was monitored over a one-week period by a fixed-wing UAV (Nathan et al., 2015). The UAV's 3.1 kg, 25 W open-path sensor had a precision of 0.1 ppmv at 1 Hz. The study confirmed the capacity of UAVs to situate themselves inside a plume of interest, and revealed important spatial and temporal variability in concentration in close-range plumes. Overall emission rate uncertainty was 55%, consisting primarily of interpolation errors, wind speed, and mixing ratios. Simultaneous MGL and on-site audit validations were consistent with UAV observations. A more recent study deployed a much smaller rotary-wing UAV, which was found to reliably detect emissions as low as 0.04 kg/h (Golston et al., 2018; Yang et al., 2018).

To date, no study to our knowledge has deployed a UAV to screen for emissions at multiple facilities during a single flight. Both fixed-wing and rotary UAVs face platform- and sensor-related limitations. The preferred sensors on other platforms are generally too large, heavy, and consume too much power to be mounted on a UAV. UAVs are therefore unlikely to contribute to M1 and M2 and may contribute to M3 if equipped with an OGI or if high-resolution three-dimensional concentration maps can be generated for individual facilities. In addition to payload and power constraints, UAVs are particularly sensitive to meteorological conditions, including high winds and rain, which not only make flying more difficult, but also make concentration measurements less reliable. Obtaining precise wind measurements with high temporal resolution is critical for generating reliable estimates of location and flux (Nathan et al., 2015). This is relatively straightforward with fixed-wing UAVs that measure airspeed with pitot instruments, but more difficult with rotary wing UAVs (Neumann et al., 2011). Low sensor frequency, when combined with high flight speeds in close proximity to a methane source, further increases uncertainty. For example, with 1 Hz sampling, a UAV had plume transects of 1-2 seconds, which is insufficient for generating an accurate spatial interpolation of concentration (Nathan et al., 2015). Regulatory limitations often restrict UAVs to a maximum flying height, within line-of-sight distance, daytime-only flying, and prohibit flying near airports or above populated areas. Finally, short flying times (typically 20 minutes to 2 hours) make UAVs unsuitable for screening in settings with low infrastructure density.

While UAVs are not yet established within LDAR, they may overcome some of the challenges they face. The development of cheap and lightweight sensors is underway (Patel, 2017), including DIAL for volatile organic compounds (Gardi et al., 2017), UAV-equipped spectrometers (Tao et al., 2015), and other high-precision sensors (Golston et al., 2017). Relaxing line-of-site regulations, improving battery life and aerodynamics, and developing solar-powered UAVs (Malaver et al., 2015; Rojas et al., 2015) may allow longer campaigns. However, for regional screening, a case for using UAVs as opposed to piloted aircraft has not been put forward.

2.5.6 Satellites

To date, satellites have only been used for M2, albeit with limited consensus on the reliability of regional attribution and temporal trends (Bruhwiler et al., 2017). The field has developed rapidly over the past two decades, and numerous new satellites have been – or will soon be – launched (Jacob et al. 2016). To date, four satellites have been used to measure methane emissions in the troposphere: SCIAMACHY, GOSAT, TROPOMI, and GHGSat, but the latter two are new and their capabilities have yet to be fully demonstrated. All these instruments use backscattered shortwave infrared (SWIR) radiation to infer column-integrated mixing ratios. Concentrations are calculated by comparing the reflected spectrum to a model spectrum, which can be generated using the ‘full-physics’ method or using CO2 as a proxy (Jacob et al., 2016). Observations can then be combined with a gridded emissions inventory in a chemical transport model, which typically uses a Bayesian inversion to produce an optimized emission field. Due to a sensor failure, SCIAMACHY only produced quality measurements between 2003 and 2005 (Frankenberg et al., 2011), with a pixel size of 30 x 60 km and 6-day repeat coverage. GOSAT, with circular pixels (10 km diameter) and 3-day temporal resolution, observes pixels only every ~260 km in cross- and along-track directions, resulting in extremely sparse data (Kuze et al., 2016). TROPOMI promises to improve on the precision and pixel resolution of GOSAT and the data continuity of SCIAMACHY at 1-day temporal resolution (Butz et al., 2012); preliminary results are promising (Hu et al., 2018).

Methane-sensing satellites outrival other technologies in spatial coverage (Schneising et al., 2014). Once operational, satellites require relatively little further investment and should collect data for many years. Column-averaged concentration measurements do not require plume access, which alleviates many of the complications typical of plume characterization. However, while methane-sensing satellites have seen promising advances in recent years, several limitations must be addressed before they can be used for facility-scale screening. Most current and prospective satellites have pixel resolutions of 4 – 100+ km2, meaning that individual pixels potentially contain multiple emitting sites and non-target sources. These satellites are unable to discriminate between thermogenic and biogenic sources, different facility types, operators, and so on, making source attribution complex or impossible. SWIR instruments require a reflective surface and are only effective on land, during the day, and in the absence of clouds (Fig. 4). Only 9% of SCIAMACHY and 17% of GOSAT retrievals are successful, leading to useful observations only once every ~67 and ~18 days, respectively (Jacob et al., 2016). Measurements in high-latitude regions may be prevented all winter by snow and low radiative flux densities. As all current satellites are in polar sun-synchronous orbit, there also exists a diurnal bias as satellites pass overhead at the same time each day, typically in the early afternoon. Finally, additional errors may arise from the chemical transport model, the bottom-up inventory, and from spatial variability in albedo due to land cover and topography (Jacob et al., 2016).

It remains unclear whether satellites will be become useful screening tools, but there is potential. The recently launched GHGSat Claire is a demonstration satellite that promises to screen for facility-scale emissions with a spatial resolution of < 50 m (Figure 2.3e), but emissions detection and measurement capabilities have yet to be publicly benchmarked. According to simulation, the detection limit for GHGSat should be 240 kg/h for winds of 5 km/h (Jacob et al., 2016), and progress is being made to improve quantification algorithms (Varon et al., 2018). This detection threshold is much higher than the average leak, encompassing the top 700 anthropogenic point sources reported in the US EPA Greenhouse Gas Reporting Program, which does not include fugitive emissions from upstream O&G (Jacob et al., 2016). The development of active LiDAR sensors operating in the SWIR is now underway (MERLIN), with plans for a joint launch in 2020 by the German Aerospace Center and the French National Center for Space Studies (Jacob et al., 2016; Kiemle et al., 2014; Riris et al., 2017). Using active sensing instruments to measure methane from space may improve spatial resolution, allow for 24-hour data collection, and reduce backscattering effects and cloud contamination. Several geostationary satellites have also been proposed to provide continuous, high-resolution monitoring over areas of interest (e.g. O&G basins), and continental coverage at hourly intervals.

2.6 The future of LDAR

2.6.1 Integrating screening technologies

Screening technologies are unlikely to replace handheld devices, which combine the sensitivity and spatial precision needed to confirm and diagnose fugitive sources. Among screening technologies, there exist notable differences in detection limits, readiness levels, flux estimation errors, spatial resolution, cost, susceptibility to adverse weather, suitability for alternative use-cases, and future potential. Thus, in selecting the appropriate sensor or platform for screening, the application must be well understood, and numerous trade-offs must be negotiated. Currently, fixed sensors, MGLs, UAVs, and aircraft have the highest readiness levels for screening, and each is suited to niche applications. Satellites, which currently have limited suitability for commercial LDAR, receive considerable investment and are likely to see ongoing innovation.

A major outstanding challenge for screening technologies is their inability to discern vented from fugitive emissions. Currently, most jurisdictions authorize facility-level venting limits, which may confound efforts to screen for fugitive emissions. False-positives during screening could mistakenly trigger follow-up surveys using close-range methods. Needless searching for these ‘ghost sources’ may increase the cost of screening and dissuade operators from moving beyond conventional LDAR. One solution is to reduce false positives by screening only for sources that greatly exceed venting limits. Should regulators impose stricter venting limits, screening techniques could become more sensitive to the relative presence of fugitive emissions. Similarly, screening could become a popular approach if regulators were to eliminate the distinction between vented and fugitive sources, opting instead for site-level limits on total emissions. Such outcome-oriented policies also provide operators the flexibility to limit emissions according to local opportunities and constraints.

Our study reveals that not enough work has been done to evaluate the performance of screening technologies. While MGLs, aircraft, and satellites have been used extensively for independent research, often with the goal of characterizing emissions from the O&G system, their suitability for use in LDAR is largely speculative. There is a pressing need for independent research to critically evaluate the strengths and limitations of screening methods, with attention to the development of detection probability curves that account for realistic environmental and operational conditions. Interest in such work is growing, good examples include the 'Mobile Monitoring Challenge' by the Environmental Defense Fund and Stanford University, and the Ginninderra experiment (Feitz et al., 2018). Ultimately, such work could lead to a policy framework for evaluating equivalence in emissions reductions among different suites of technologies and methods. Furthermore, such efforts can filter through the excitement typically associated with new technologies and provide comprehensive and reliable assessments of their capabilities. However, there are additional hurdles to overcome before new technologies are assimilated. These include the perceived unreliability of new technologies, lack of familiarity with underlying principles of measurement, requisite changes to field and data management practice, and shifts in regulatory language (CATF, 2013). To overcome these challenges, regulators should develop a decision-making process to test and approve new technologies.

This review suggests that effective LDAR solutions could leverage different technologies based on their context-dependent strengths and limitations. A promising approach to cost-effective LDAR is a comprehensive monitoring program (CMP) that integrates two or more screening and close-range technologies. To date, several studies have explored the use of multiple platforms for methane detection and quantification. For example, the Barnett Coordinated Campaign of 2013 (Harriss et al., 2015; Zavala-Araiza et al., 2015) consisted of numerous measurements from manned aircraft (Karion et al., 2015; Lavoie et al., 2015), MGLs (Rella et al., 2015; Yacovitch et al., 2015), and a UAV (Nathan et al., 2015), synchronized over time and space. Many other examples of integrated measurement campaigns can be found in the literature (Babilotte et al., 2010; Bateman et al., 2016; Frankenberg et al., 2016; Gardi et al., 2017; Gemerek et al., 2017; Kort et al., 2014; Schwietzke et al., 2018; Subramanian et al., 2015). However, most of these integrative campaigns were designed to improve methane emissions estimates without practical consideration for use in LDAR programs. Studies that integrate technologies in different configurations to optimize both cost and mitigation are needed, and they must not be blind to the practical considerations faced by operators. To date, only one study has compared a CMP to conventional LDAR (Schwietzke et al., 2018). Here, the authors found that using aircraft to direct ground surveys could be at least as cost-effective as conventional LDAR; compared to ground teams, the aerial surveys detected up to 26 times more methane from half as many sources. However, these results can be highly context-dependent, and the work illustrates the confounding influence of vented emissions on CMP effectiveness.

In areas with multiple interspersed companies or with low infrastructure density, screening by means of aircraft and satellite deployment may not be economically viable unless costs are shared. Operators and regulators should therefore explore different costing scenarios for screening. For example, satellites and piloted aircraft could be used for regular large-scale surveillance, financed using a subscription-based approach, and operated as part of a government program or by a third party. Regular UAV campaigns over high-density and/or high-risk areas could supplement these data, providing quick estimates of high-priority sources needing immediate attention. Finally, MGLs equipped with handheld devices could use intelligence from large-scale campaigns to determine where and when to investigate further or conduct repairs. A simpler CMP could integrate a handheld approach (Method 21 or AWP) with a single screening technology.

These general examples of possible CMP configurations are a simplified abstraction of what would be a multifaceted undertaking riddled with technological, logistical, and regulatory challenges. As technologies and methods evolve, so will the most appropriate CMP configuration for a given application. Some technologies may see considerable improvement in capabilities, while others may become obsolete. UAVs and satellites, as relatively young platforms, are particularly well-suited to overcome the limitations that currently prevent them from playing a larger role. As sensors become smaller, and UAV flight times increase, the UAV niche may grow. Full UAV autonomy for this application only requires leveraging existing robotics technology but will take time to mature. Similarly, if an increasing number of methane-sensing satellites are deployed, and if their capacity to deliver higher-resolution data improves, satellites could supplant manned aircraft for regional surveys. Passive, continuous measurement may also become prominent as sensors become less expensive and more durable, which could reduce MGL labour costs as sensing is accomplished on vehicles performing other tasks. Ultimately, the most popular programs will achieve compliance at the lowest cost. Monitoring plans, largely dictated by economics, may guide LDAR away from labour-intensive and towards automated methods. As new technologies are approved for LDAR, competition could increase innovation and reduce monitoring costs. This could lead to a win-win situation for the public and industry, as an increasingly greater proportion of the leak-size distribution can be repaired at a net-positive cost. Given the range of possible screening scenarios and technology metrics that must be considered, providing explicit guidance on when each technology should be used and in what combination is beyond the scope of this article. Ultimately, models should be developed to evaluate the most effective CMP for a given context, especially if multiple technologies are to be used. These models could consider infrastructure density, monitoring cost, detection limits, meteorology, and other factors. At present, these models are informed by a limited empirical understanding of available technologies (Kemp et al., 2016).

2.6.2 Beyond technology integration

Future CMPs could go beyond integrating close-range and screening technologies. The collection, management, analysis, and distribution of emissions data would contribute greatly to the success of targeted campaigns. Despite considerable efforts to develop robust emissions factors (Allen et al., 2014b, 2014a, 2013; Johnson et al., 2015; Kang et al., 2014; Littlefield et al., 2017; Marchese et al., 2015; Michanowicz et al., 2017; Omara et al., 2016), much more data are needed to investigate predictive analytics based on assumed risk factors (e.g. management, age of infrastructure, geology, etc.) Some have already begun to make progress in this area, introducing opportunistic mobile sensing using meteorological data as well as facility-specific information such as age and production rate (Albertson et al., 2016). Others have worked to integrate greenhouse gas measurements from various platforms with data products that estimate urban emissions (Davis et al., 2017). Despite being one of the most important questions for modelling and understanding methane emissions from O&G, we still have a limited understanding of the nature of the emission-size distribution of different regions or facility types and ages. While it is established that these distributions are heavy-tailed (Brandt et al., 2016), we are only beginning to understand the causes of super-emitting sources (Zavala-Araiza et al., 2017). An understanding of the temporal variability of emissions also remains elusive, although most evidence suggests that emissions are often intermittent (Allen et al., 2017; Englander et al., 2018; Schwietzke et al., 2018). Multi-platform comparisons of not just the same region, but of the same plumes, would lead to an improved understanding of temporal variability and could greatly improve localization and quantification capabilities. Emissions data from facility operators can be difficult to obtain, as standardized collection and reporting for event-specific fugitive emissions has not yet been mandated, adding further incentive to developing CMPs with data-driven prediction capabilities.

Improving incentivization structures at all levels of industry would help to further ensure that technologies are being used to their potential. Principal-agent problems must be identified and addressed. For example, those who own the infrastructure – especially in the case of pipelines – do not necessarily own the leaking product. Furthermore, an improved understanding of how human error influences detection likelihoods for different technologies is needed. Handheld cameras are especially problematic in this regard, due to the subjective nature of the instrument operation and data interpretation (Footer, 2015; von Fischer et al., 2016). Incentivizing LDAR among operators might be achieved by improving our understanding of the economic benefits (e.g. product loss prevention) of investing in mitigation (Palacios et al., 2017). In the future, carbon pricing of fugitive emissions could also be used to incentivize mitigation.

2.6.3 Regulatory and economic considerations

Clear, predictable, and enforceable policies are needed to ensure the effective implementation of LDAR programs. Although the financial return on LDAR programs is often modelled to be net positive (Kemp et al., 2016), as lost gas can be sold to offset LDAR costs, operators may have more pressing or promising investments to make, or insufficient capital to invest in new technology. While the economics vary by jurisdiction, we identify three broad classes of incentivization, similar to Ravikumar and Brandt (2017): 1) Direct regulatory forcing, in which operators are forced to comply with regulations; LDAR is regulated in detail, and compliance is assessed against standards of operation; 2) Indirect regulatory forcing, whereby emissions are taxed to offset associated externalities; 3) Voluntary mitigation programs, where companies design and implement the LDAR program they find has the best return on investment, while complying with health and safety regulations.

The type of LDAR implementation framework has important implications for the effectiveness, quality, and capacity of these programs to mitigate emissions. With direct regulatory forcing (class 1), the regulator must be able to implement and enforce suitable protocols. Compliance is assessed according to the standards set by the regulator, not necessarily the absolute reduction in GHG emissions. New technology is difficult to implement, as there is a delay between establishing appropriate standards of operation and achieving regulatory approval. With indirect regulatory forcing (class 2), a reliable estimate of the actual GHG emissions must be acquired to achieve compliance. Class 2 has only recently become a consideration, as methods for quantifying emissions have matured (Atherton et al., 2017; Yacovitch et al., 2015). The way operators meet compliance is open, and there is impetus for innovation, as there is a competitive market for improving methods and technology. Without regulations (class 3), fugitive emissions may rise if infrastructure is neglected due to limited capital or more important alternative investment opportunities.

The success of a direct regulatory model is limited by unknowns, such as the mitigation effectiveness of different LDAR programs. Sustained long-term measurements at natural gas facilities could help to inform future mitigation policies. By measuring pre- and post-LDAR emissions factors through different implementation periods (e.g. 0, 3, 6, and 12 months), we could better constrain long-term emissions trends. Such efforts may also help identify the sources that are prone to relapse following repair. At the policy level, this can translate into more directed regulations – equipment that doesn't emit after repairs can be inspected at a much lower survey frequency, thereby reducing costs. As hazardous air pollutants are often co-emitted with methane, opportunities exist to improve efficiency by harmonizing monitoring efforts.

2.7 Conclusion

Current LDAR programs rely on close-range methods such as Method 21 and AWP. While close-range instruments are indispensable for identifying and documenting component-level fugitive sources, they are relatively labour intensive. Rather than relying exclusively on handheld instruments, regulations in Canada and the US are moving towards the integration of screening technologies. Given the characteristic shape of most leak-size distributions, frequent screening for super-emitters could reduce fugitive emissions and lead to a lower, albeit more targeted reliance on exhaustive close-range surveys. Fixed sensors, MGLs, UAVs, manned aircraft, and satellites, have been used for research-based applications and for monitoring other air pollutants, but are only just gaining interest as tools for LDAR. As screening technologies, each of these technologies is uniquely suited to a range of environmental, economic, and operational contexts. Fixed sensors, MGLs, UAVs, and aircraft are arguably ready for integration as screening products into current LDAR programs. Satellites may soon be ready given anticipated innovation.

To meet emissions reduction targets and reduce monitoring costs, governments and the O&G industry should consider CMPs that integrate different technologies into a multi-scale, data-driven, methane-sensing system. CMPs could be tiered both spatially and temporally, with frequent monitoring at coarse spatial resolutions using screening technologies, and infrequent, targeted monitoring at fine spatial scales using close-range methods. In addition to cost-effective monitoring and enhanced mitigation, CMPs could improve scientific understanding of how, when, and why fugitive emissions occur, and enable dynamic inventories, regulatory accounting, and evaluation of mitigation success. Cooperation and transparency among regulators, O&G companies, monitoring agencies, and researchers will be crucial for moving towards a CMP model. Regulatory flexibility and standardized protocols for the approval of new technologies must be developed. Finally, there is a need for research evaluating individual technologies and CMP configurations for their mitigation potential, economic viability, and regulatory compliance.

Chapter 3

A methane emissions reduction equivalence framework for alternative leak detection and repair programs

3.1 Abstract

Fugitive methane emissions from the oil and gas sector are typically addressed through periodic leak detection and repair surveys. These surveys, conducted manually using handheld leak detection technologies, are time consuming. To improve the speed and cost-effectiveness of leak detection, technology developers are introducing innovative solutions using mobile platforms, close-range portable systems, and permanent installations. Many of these new approaches promise faster, cheaper, or more effective leak detection than conventional methods. However, ensuring mitigation targets are achieved requires demonstrating that alternative approaches are at least as effective in reducing emissions as current approaches – a concept known as emissions reduction equivalence. Here, we propose a five-stage framework for demonstrating equivalence that combines controlled testing, simulation modeling, and field trials. The framework was developed in consultation with operators, regulators, academics, solution providers, consultants, and non-profit groups from Canada and the U.S. We present the equivalence framework and discuss challenges to implementation.

3.2 Introduction

The International Panel on Climate Change (IPCC) recently underscored the importance of reducing methane emissions to keep global warming below 1.5 °C (IPCC, 2018). However, global natural gas production is increasing, with an estimated 1.7 – 2.3% of total production (primarily methane) escaping directly to the atmosphere (Alvarez et al., 2018; International Energy Agency, 2017). Leak detection and repair (LDAR) programs are the most common regulatory tool for mitigating fugitive methane emissions (leaks) from upstream oil and gas. Historically, LDAR programs have relied on a variety of close-range methods implemented through U.S. Environmental Protection Agency's (EPA) Method-21 or Alternative Work Practice to perform component-level surveys. Although effective, these approved methods remain labour-intensive (ICF International, 2015, 2014). Recently, new methane-sensing technologies have emerged, promising faster, cheaper, or more effective leak detection (Fox et al., 2019a). In response, regulators in Canada and the U.S. have created opportunities for flexible LDAR programs that permit new approaches to detection. However, operators wanting to move from a 'regulatory' to an 'alternative' LDAR program are typically required to demonstrate equivalence in emissions mitigation (Government of Canada, 2018). Regulatory approval of new technologies will be effective only if there is a transparent framework for operators and solution providers to demonstrate equivalence.

A broad spectrum of candidate technologies exists for potential integration with alternative LDAR programs. The most common alternative technology classes include handheld instruments, mobile ground labs (Caulton et al., 2018; Yacovitch et al., 2015), unmanned aerial vehicles (Barchyn et al., 2017; Golston et al., 2018; Nathan et al., 2015), stationary sensors (Coburn et al., 2018), manned aircraft (Conley et al., 2016; Frankenberg et al., 2016; Terry et al., 2017), and satellites (Jacob et al., 2016). Although initiatives such as the Environmental Defence Fund / Stanford Mobile Monitoring Challenge and the ARPA-E MONITOR program have sought to evaluate some of these technologies, little progress has been made in systematically comparing alternative and conventional LDAR programs. Here, we propose a framework for demonstrating emissions reduction equivalence between LDAR programs.

3.3 Framework development

The framework was originally developed at a multi-stakeholder workshop and has since been publicly reviewed at two additional workshops and a 30-day public comment period. The development process was designed to be transparent and inclusive of diverse opinions and interests. Each of the workshops followed Chatham House rules.

On 25 July 2018, approximately 50 scientists, regulators, operators, consultants, and non-profit organizations gathered at the University of Calgary in Alberta, Canada, to discuss and solicit perspectives on how to demonstrate equivalence for alternative leak detection and repair (LDAR) programs (Figure 3.1). The workshop was organized into three sets of presentations followed by break-out sessions. Presentations were used to establish a common understanding of the equivalence challenge, including regulatory context, industry needs, and scientific knowledge. During break-out sessions, mixed stakeholder groups of 8-10 participants engaged in semi-structured discussions around three themes: (a) Thinking about equivalence, (b) Developing a common framework, and (c) Applying equivalence to specific technologies. The following day, a committee of 8 experts met to distill these conversations into a draft framework. The resulting white paper was publicly distributed, and comments were solicited for 30 days to enable contributions from those unable to attend the workshop.

Figure 3.1 Number of workshop participants by stakeholder group.

On 8 January 2019, a second two-day workshop was held at Colorado State University in Fort Collins. The draft framework was presented to 68 stakeholders from Canada and the U.S., proposed amendments were debated, two modifications were made, and the framework was collectively approved. A third workshop was held on 14 February 2019 geared specifically towards LDAR solution providers; this group was not invited to earlier workshops to prevent bias and manage workshop attendance. Participants at Workshop 3 were supportive of the framework and no further modifications were proposed.

3.4 Definitions

In the context of equivalence, it is important to distinguish among technologies, methods, and programs:

> A **technology** is a gas sensing instrument, optionally configured with a deployment platform and/or ancillary instruments (e.g. anemometers, positioning), that can be used to gather data on emissions.

A **method** combines a technology, a work practice, and analytics for use in an LDAR program. A method must clearly state any mandatory actions to be performed as part of the work practice, along with suitable operating conditions for the technology. These can include environmental conditions, limitations on facility-types, technology configurations, and survey procedure.

An **LDAR program** is the systematic implementation of one or more methods across a collection of assets. The program describes the method, or combination of methods, to be used for each facility, along with survey frequency, repair response, and reporting standards. Ultimately, it is the LDAR program that results in emissions mitigation, not the technologies or methods in isolation.

The frequently used term ‘technology equivalence’ is a misnomer, as no two technologies can be shown to have equivalent mitigation potential outside the context of a method and/or program. Although mitigation equivalence may be demonstrated among methods, it is most universally demonstrated at the program level for three reasons. First, multiple methods may be used simultaneously in a program. Assessing equivalence for multi-method programs is not as simple as aggregating the mitigation from individual methods due to potential detection overlap. Second, mitigation is a function of survey frequency, which is typically part of the program, not the method. For example, EPA’s Method 21 is a method that can be implemented at different frequencies (e.g. monthly, quarterly, annually) as part of a program to achieve different targets. Third, depending on regulatory language, ‘alternative LDAR’ doesn't necessarily require adoption of new technology. Operators may want to use approved methods while otherwise adjusting the program (e.g. definition of a leak, type of equipment surveyed, repair requirements, survey frequency, etc.) Operators may even propose different survey protocols for different asset types or locations. Despite using existing methods, these alternative programs may also need to demonstrate equivalence.

3.5 Equivalence framework

We define an equivalence framework as a scientifically-rigorous and transparent process that uses a combination of empirical data and modeling to estimate emission reductions from the implementation of an LDAR program and compares this estimate to mitigation from an approved program or a defined target. The reference mitigation achieved by the approved program and the spatial scale of comparison must be specified by the regulator. The proposed framework was designed to be of general interest to regulators developing alternative LDAR policy for conventional and unconventional oil and gas regions and does not account for specific jurisdictional contexts. It consists of five stages (Figure 3.2):

1. Method identification to assemble and define new methods
2. Controlled testing to evaluate the performance of new methods
3. Simulation modeling to predict the performance of new programs
4. Field trials to establish operational efficacy of new programs
5. Full approval of the alternative LDAR program

Stages 1 and 2 focus on methods while subsequent stages require a program. The five stages will require engagement from multiple stakeholders including solution providers, operators, independent evaluators, and regulators. Stakeholders may wish to use the results of one stage to inform progression through the framework. An adaptive feedback process would help transfer experience and knowledge among stages.

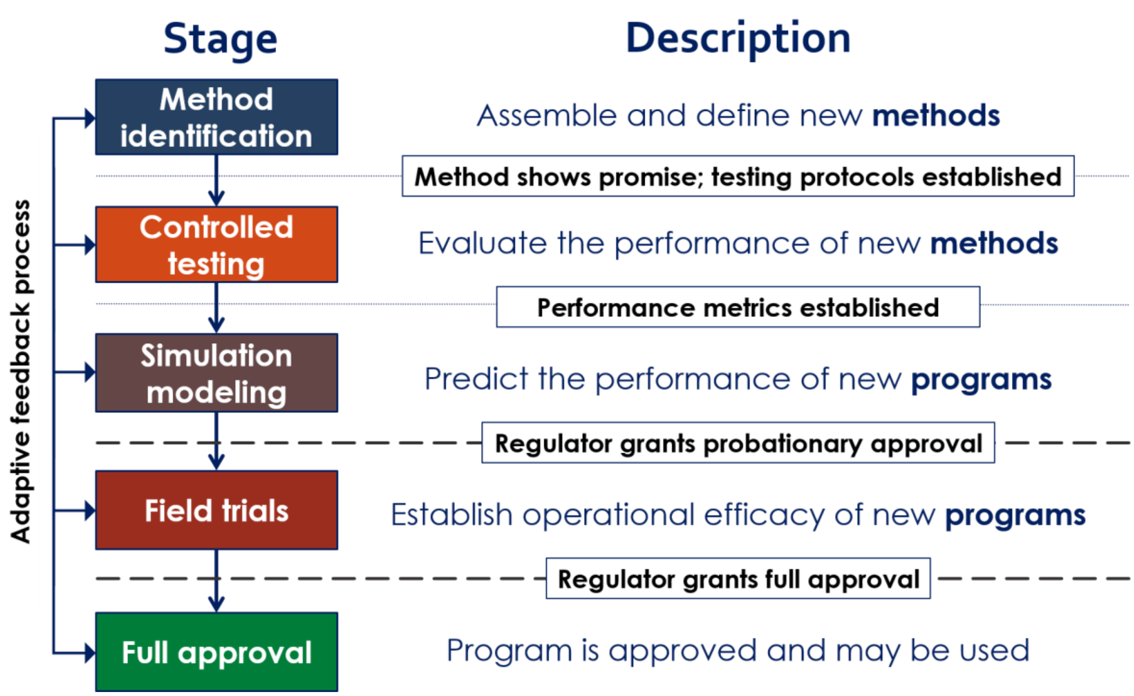

Figure 3.2 Overview of the equivalence framework. Stages are separated by objectives, whether informally (fine line) or by regulatory approval (dashed line). An adaptive feedback process connects each of the five stages to enable sharing of knowledge and evolution of the demonstration process.

3.5.1 Stage 1: Method identification

Clear method identification is critical to developing effective protocols for testing and evaluation. During Stage 1, applications should be solicited for new methods seeking to demonstrate equivalence. Clusters of similar methods can then be organized and defined to (a) identify common features and constraints, and (b) establish protocols for controlled testing. Specifically, Stage 1 will set group-wise performance metrics that link controlled testing with modeling. Standardized evaluation of multiple methods could improve comparisons, especially during controlled testing, and enable synchronized controlled testing to reduce costs. If performance of approved standards is unknown, these methods should also be moved through the Stages 2-4 to establish a reference mitigation target.

3.5.2 Stage 2: Controlled testing

Standardized controlled testing is necessary to understand method performance. Single-blind controlled field testing, administered by independent experts, should be used to develop performance metrics across a range of operational conditions. In addition to generating performance metrics, controlled field testing can contribute to the development and promulgation of clear and reproducible testing standards. Development of testing protocols should be led by neutral experts supported with commenting from the broader user, developer, and regulatory community.

Performance metrics should be carefully selected to ensure valid empirical inputs for the subsequent modeling stage. For each method, detection probabilities should be established under a range of conditions. Additional metrics should also be developed, depending on the method, and could include quantification accuracy, false positive rates, and spatial resolution. A geographically-dispersed network of test-sites could facilitate testing under different environmental conditions and facility types. Member sites should collectively adhere to testing and reporting standards recognized by all stakeholder groups. Joint funding of this network by stakeholders standing to benefit from emerging methods could minimize financial burden on solution providers, while maintaining independence of the testing. In addition to testing results, technology developers would also benefit from a testing process that identifies ways to improve their technology, work practice, and analytics. At this early stage, operators may not need to be involved, but may wish to partner with solution providers to develop methods that fit their needs.

3.5.3 Stage 3: Simulation modeling

Given the highly-skewed and stochastic nature of fugitive emissions from upstream oil and gas (Brandt et al., 2016), fully-representative testing of methods and programs would be prohibitively expensive. Simulation modeling is a fast and cost-effective way to evaluate and explore a range of possible LDAR program configurations, forecast performance over long periods, and develop programs with cost- or mitigation-optimized deployment of different methods across a collection of assets. Stage 3 would use the performance metrics developed in Stage 2 as inputs for simulating the mitigation effectiveness of an alternative program. Whereas Stage 2 evaluates LDAR methods for detection effectiveness, Stage 3 evaluates LDAR programs for aggregate mitigation impact. The simulations will estimate total emissions detection over a reporting period, which can translate into emissions reductions when repairs are conducted.

Controlled testing standards, protocols, and outputs must be designed with modeling needs in mind, and models should strive to represent new technologies accurately. To date, modeling tools such as FEAST have been used to compare costs and mitigation effectiveness of LDAR programs (Kemp et al., 2016). Extending such tools to demonstrate LDAR program equivalence may require new functionality. Simulation models must balance fidelity, data input requirements, accessibility, and resilience to gaming. Developers of prospective methods should be encouraged to experiment with models before entering Stage 1 to estimate performance requirements and avoid committing resources to testing if not ready. By the end of Stage 3, technology developers could decide whether to adjust their method and perform more testing before moving forward in the framework. In most cases, modeling results would need to show improved emissions mitigation and/or reduced costs to be attractive to operators, who are typically responsible for submitting alternative LDAR program applications. Regulators can help guide the development of modeling and reporting requirements. Jurisdictions should provide representative inputs (e.g. baseline emissions, leak-size distributions, and activity factors) that best reflect assets under their control, because high-quality empirical inputs generate more accurate outputs that can improve decision-making capacity.

The modeling stage is focused primarily on emissions mitigation and is unlikely to incorporate cost. For a proposed program to be attractive, operators and solution providers will need to develop cost models in parallel to determine whether the program warrants further work.

3.5.4 Stage 4: Field trials

Controlled testing and modeling may fail to capture the full scope of real-world performance, including unforeseen human and environmental factors that may only become apparent during deployment. Data from the field may also help improve models and build confidence in their predictions. We therefore recommend field trials to evaluate performance in operational conditions and demonstrate the efficacy of candidate programs.

To initiate Stage 4, operators and technology developers would work together to develop an alternative LDAR program application for submission to the regulator. The submission would include results from Stages 2 and 3, methods to be used, survey frequencies, reporting protocols, a field trial plan, and other information relevant to the alternative program. The duration of the field trial (in time or number of surveys) and the number of assets involved should be specified in the application to ensure a representative sample size. The regulator would review the application package and, if satisfactory, may approve the proposed field trial.

During Stage 4 the alternative program would be implemented on a specified proportion of the operator's assets. Evaluating field trial effectiveness could take several forms. At minimum, a brief field trial should be required to troubleshoot for unanticipated issues not accounted for or predicted by modeling. The bigger goal – evaluating mitigation effectiveness in the field – is complicated because (a) true emission rates are unknown, (b) most quantification methods are highly uncertain, and (c) obtaining a representative sample including 'super-emitters' – a small number of high-emitting sources – may be cost-prohibitive. Limited insights into mitigation effectiveness may be gleaned by deploying alternative technologies alongside conventional LDAR and a range of component- to facility-scale quantification techniques. Selecting a field trial approach may depend on jurisdictional policy context. For example, Canadian regulatory language considers operators under compliance during field trials. However, in other jurisdictions, a full regulatory LDAR program must be implemented in conjunction with field trials.

3.5.5 Stage 5: Full approval

Results from all stages would be communicated to the regulator by the applicant for evaluation and full approval. Program auditing and compliance details may be included in the application, along with the program scope. Upon full approval, operators could substitute the new method for existing (approved) LDAR methods. Regulators should consider developing commensurate frameworks for approvals, allowing alternative programs approved in one jurisdiction to be more easily approved in another, with necessary adjustments for geographical constraints, gas compositions, weather, and other relevant factors.

3.5.6 Exceptions

Should a regulator choose to adopt a version of this framework, they may wish to identify situations in which exceptions are warranted. We provide three examples, but others likely exist:

1. Operators may wish to implement an alternative LDAR program using approved methods. In such a case, Stages 1-2 of the framework could be skipped.
2. A novel method may be proposed that is very similar to an approved method. In such a case, a reduced suite of controlled testing scenarios may be warranted.
3. A company may want to implement an alternative practice specific to an individual facility or type of facility. In such a case, it may be reasonable to skip directly to field trials.

3.6 Challenges

Several unresolved challenges must be considered before adopting this framework. We arrange these challenges into five broad categories: controlled testing, modeling, scale and source disambiguation, human factors, and logistics.

3.6.1 Controlled testing

Each method has a unique set of environmental and operational variables that influences its performance. These variables must be identified and incorporated into testing, and results must be used to inform when, where, and how new methods should be deployed. Two distinct issues can arise from omitting critical variables from testing. First, a method could be tested beyond its optimal use case. As a hypothetical example, a method may only work at 10% of sites, but detect 90% of emissions at those sites. Without context, the method could be defined to detect only 9% of emissions. Similar consideration may be required for meteorology, daylight, topography, work practice, or other method-specific factors. Second, if performance metrics are developed under optimal measurement conditions, a method may fail to achieve anticipated mitigation when deployed. A potential solution is to develop ‘operational envelopes’ that define the measurement conditions under which a method can be deployed. Operational envelopes would reflect testing conditions and could expand over time with additional testing to access new markets. Criteria for expanding operational envelopes must balance cost and scope of testing; fully representative testing under all environmental conditions is not practical.

3.6.2 Modeling

Critical modeling challenges include establishing baseline emissions distributions, specifying functionality, and model validation. Emission rates in leak simulation models are sampled from empirical distributions. Several distributions exist, but they differ by basin and represent only a snapshot in time. For most producing regions, data are incomplete, dated, or do not exist. Acquisition of detailed baseline measurements is time-consuming, expensive, and can suffer from measurement bias or uncertainty. Sample sizes must be high to account for super-emitters; recent studies have demonstrated that the top 5% of sources typically contribute 50% of total emissions (Brandt et al., 2016). Chance variability in the presence and magnitude of super-emitters can therefore result in markedly different distributions, which may favour certain methods over others during modeling.

Optimizing model functionality and accessibility is an unresolved challenge. On one hand, modeling tools that are intuitive, accessible, and transparent will be widely used and accepted. On the other hand, these tools must be able to accommodate a broad diversity of methods, environments, and policy contexts. With insufficient functionality, methods could be excluded or poorly represented. Finally, model validation is difficult. While detection modules for individual methods may be validated in the field, validation of long-term programs with multiple methods will be challenging.

3.6.3 Scale and source disambiguation

Method 21 and handheld cameras are well-established in LDAR as they enable component-level detection, which can often lead to immediate diagnosis and repair. However, many emerging methods propose rapid screening for aggregate emissions, often at the facility scale. If anomalous emissions are detected while screening, close-range methods must be deployed to confirm and diagnose the source. Screening-based programs must therefore articulate when and how repair teams will respond to detection events. However, screening is sensitive to confounding sources. For example, most sites have legal venting limits that can be difficult to distinguish from fugitive emissions. The possible presence of confounding sources may increase the rate of false-positives, leading to unnecessary follow-up surveys.

3.6.4 Human factors

Human dimensions of LDAR have not been studied. In the context of equivalence, the biggest challenges are system gaming and post-approval incentivization. If the framework is recognized, it will become a barrier to market access for emerging solution providers, who will face pressure to succeed. Simulation models must be protected from directed attacks, such as modifying source code to improve results or selectively editing input data sets. Less nefarious deception, such as selective reporting of data or results, could occur. These and other temptations must be thoroughly considered, and preventative measures implemented. Model results should also be reviewed for accidental misapplication (e.g. inappropriate selection of input data).

The framework faces two broad incentivization challenges: program improvement and program compliance. First, solution providers may be disincentivized to improve work practices, technologies, and analytics once they are approved if modifications risk voiding approval. To prevent stagnation, efficient approval mechanisms for updates should be implemented. Second, service providers must be incentivized to abide by approved programs as defined, particularly when human intervention is required to complete the approved process. For example, technicians may face pressures from employers, operators, or unpleasant working conditions (e.g., excessive cold or heat) that may decrease program effectiveness relative to the approved program. Performance may also vary with user experience, and method-specific training protocols should be defined and implemented.

3.6.5 Logistics

Several logistical challenges remain unresolved. Stakeholder roles and responsibilities must be defined, including funding sources, oversight of controlled testing, and development, management and administration of the simulation model. Transparency of the demonstration process must be established, including whether performance results are made public, standards for protection of intellectual property, and whether approved methods and programs are made available for all operators. These challenges may be resolved according to jurisdictional differences in regulatory language. Jurisdictions should work together to ensure that definitions and approval standards are consistent to minimize redundant bureaucratic barriers to approval.

3.7 Conclusions

This framework is a first step towards encouraging adoption of alternative LDAR programs. Implementation should strive to balance rigor and practicality. If hurdles are too great, operators will avoid alternative methods and settle for regulatory LDAR, which may lower investment, stifle innovation, and limit our ability to reduce emissions and learn about new methods through deployment. However, rigor in framework implementation is necessary, as failure to curb fugitive emissions may not be evident due to the challenge of tracking baseline emissions.

Moving forward, work will be required to consolidate, refine, and execute the framework. First, relevant stakeholders should be identified, and their roles clearly defined. Regulators may want to take responsibility for leading the development of a formalized framework with clear and detailed criteria for demonstration. A collaborative network of controlled testing sites should be developed, with broad geographical representation, reliable funding, and independent operation. New methods must be formalized, and testing protocols developed. Open-source simulation models should be developed to be flexible, transparent, robust, and accessible. Communication networks among regulators should be established to facilitate inter-jurisdictional translatability of methods and programs. Finally, the framework should evolve to have specific guidelines for each stage, providing all stakeholders a clear understanding of the resources required to develop and implement new methods and programs.

Chapter 4

An agent-based model for estimating emissions reduction equivalence among leak detection and repair programs

4.1 Abstract

Alternative leak detection and repair (alt-LDAR) programs are being introduced by regulators to provide flexibility in how oil and gas producers manage fugitive methane emissions. However, emissions reduction equivalence must be established between a proposed program and a regulatory standard. We present LDAR-Sim, an open-source, agent-based numerical model for evaluating LDAR program performance. Novel functionality includes the ability to: (1) set facility-specific LDAR requirements, (2) simultaneously deploy multiple technologies, (3) integrate screening and close-range methods, (4) include environmental constraints, and (5) explore the impact of design emissions on screening technologies. We demonstrate LDAR-Sim in two case studies with diverse LDAR programs deployed on real assets. We find that equivalence determinations depend on explicit definition of reference standards, including weather and labour availability. Screening program performance is sensitive to the confounding presence of design emissions and to decision-making strategies that guide follow-up inspections. When comparing programs, differences in simulated performance are sensitive to leak production and null repair rates, two elusive parameters used in previous studies. Future work should better constrain empirical inputs and validate specific LDAR programs as alt-LDAR deployment increases. Regulators and producers can use LDAR-Sim to facilitate the adoption of emerging technologies but should be aware of its limitations.

4.2 Introduction

Oil and gas systems are the largest source of anthropogenic methane emissions in Canada and the US (EPA, 2019; Government of Canada, 2016). In the upstream O&G sector, methane can be emitted by design (i.e. intentional vented emissions) or as accidental leaks (i.e. unintentional fugitive emissions). To address fugitive emissions, regulators often require producers to implement leak detection and repair (LDAR) programs (AER, 2018; Government of Canada, 2018). Most LDAR programs consist of periodic surveys using handheld instruments such as optical gas imaging (OGI) cameras or organic vapour analyzers. These ‘conventional’ surveys are conducted at the component scale using close-range methods with high sensitivity.

Alternative approaches to LDAR (alt-LDAR) may exist (Fox et al., 2019a; Schwietzke et al., 2018). We define alt-LDAR as any deviation from a prescribed regulatory LDAR program (reg-LDAR). Over the past decade, a variety of technologies, methods, and program designs have emerged that may contribute to improving the cost- and mitigation-effectiveness of LDAR. In response, various regulatory agencies in the US and Canada have moved to recognize alt-LDAR as a potentially viable path to compliance (AER, 2018; CDPHE, 2018). Regulators often require evidence that proposed alt-LDAR programs achieve equivalent mitigation to the reg-LDAR standard. A combination of controlled testing, simulation modeling, and field trials has been recommended (Fox et al., 2019b).

Simulation modeling provides several advantages as a method of understanding the emissions reductions potential of an LDAR program. First, simulations allow a long-term estimation of emissions dynamics. Second, direct testing of LDAR tools on real leaks is slow, expensive, and possibly unfair as a representative distribution of emission rates is difficult to acquire (Brandt et al., 2016). Third, empirically measuring total mitigation progress is challenging due to the high uncertainties typical of most quantification methods (Ravikumar et al., 2019). The first and only existing peer-reviewed LDAR simulation tool was the Fugitive Emissions Abatement Simulation Testbed (FEAST; Kemp et al., 2016). FEAST generates a virtual leak field in which leaks arise and are removed stochastically at the component scale following a Markov process. Detection modules are then applied to the leak field. When a survey is conducted at a component, detection is evaluated by simulating a Gaussian plume and considering the physical properties of the sensor. If detected, leaks are repaired and the component's Markov state switches to not emitting. Mitigation is calculated as the difference between time-integrated emissions under the LDAR program and a 'null' scenario that estimates baseline emissions in the absence of LDAR.

The original version of FEAST is a valuable tool but has drawbacks that may limit its broader utility. First, the success of an LDAR program depends not only upon the theoretical detection capabilities of a method, but also upon practical considerations that vary by producer, facility, environment, and policy context. For example, an aircraft-based remote sensing system may perform poorly in high-latitude winters due to snow cover, cloud cover, and low solar irradiance. An alt-LDAR program that commits to monthly surveys using this system may only be successful in specific geographical or seasonal contexts. If only a handful of days are suitable for deployment, labour and equipment availability may become limited, especially if only one aircraft or crew exists. Second, LDAR simulation tools should provide producers with an opportunity to explore and optimize across a range of possible programs. For example, a producer may propose a program that simultaneously integrates aircraft, trucks, and OGIs with facility-specific survey schedules and frequencies, and custom repair protocols. Important interaction effects may emerge from concurrent deployments that could influence the equivalence determination.

To meet these needs, we present LDAR-Sim, an open-source agent-based framework that permits the simultaneous deployment of multiple methods, each with multiple agents (crews), while allowing for the definition of custom LDAR programs. We first present a detailed description of LDAR-Sim, then demonstrate the model with a case study using real assets in Alberta, Canada. We conclude with a comprehensive sensitivity analysis (SA) and discussion of modeling assumptions and shortcomings, and their implications for generating equivalence determinations.

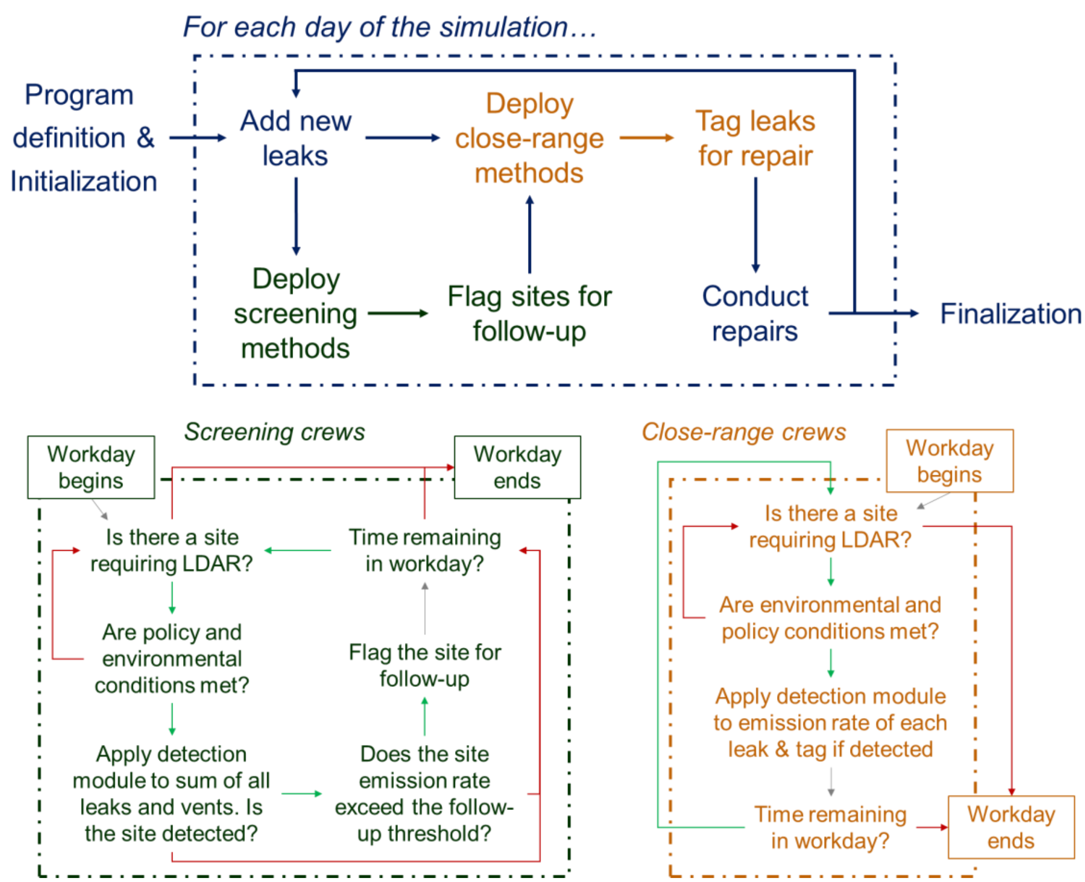

Figure 4.1 Overview of LDAR-Sim. Green text is used for screening methods and orange text for close-range methods. Red arrows represent ‘no’, green arrows are ‘yes’, and grey arrows are mandatory.

4.3 Model description

LDAR-Sim is written in Python 3 and uses the *initialize-update-finalize* Basic Model Interface (BMI) of the Community Surface Dynamics Modeling System (Peckham et al., 2013). The BMI provides a standardized, modular, and intuitive model structure to facilitate collaboration, accelerate learning and streamline coupling with external software. In LDAR-Sim, each program consists of a set of assets, one or more leak detection methods, and a set of rules governing deployment and repair. Assets are combined with empirical inputs to initialize a leak field. A dynamic simulation with daily timesteps and minute-scale temporal resolution is then used to evaluate the program (Figure 4.1). Leaks arise stochastically and are detected probabilistically according to method-specific empirical detection data. Aggregate-scale detections (i.e. by screening methods) can produce flags, which trigger follow-up by close-range methods. Leaks detected at close-range are tagged and enter a repair queue. Tagged leaks are removed from the simulation following user-defined reporting delays and repair intervals. Parameterization, selection of empirical inputs, and sensitivity analyses (SA) should be done on a case-by-case basis; no defaults exist.

In this paper, LDAR-Sim is presented in the context of demonstrating emissions reduction equivalence among different LDAR programs. However, LDAR-Sim could be used by various stakeholders for other applications. For example, policymakers could use LDAR-Sim to estimate whether existing or proposed regulations will achieve emissions reduction targets. Regulators might use LDAR-Sim to understand how to deploy different technologies as part of compliance efforts. Technology innovators might also use LDAR-Sim to understand how their solutions compare to existing technologies, to identify ideal environments for deployment, or establish innovation trajectories (e.g., sensor sensitivity requirements).

4.3.1 Program definition

A primary motivation in developing LDAR-Sim is to enable the precise definition of custom LDAR programs. Defining a program requires identifying the detection method(s), the asset base, and deployment protocols. In practice, each method module should be designed in accordance with the technology and work practice proposed by the solution provider. Ideally, a set of relevant performance metrics are developed through controlled testing. We provide a set of general method templates, including OGI, a mobile ground laboratory (MGL), and an aircraft system. Each method can be parameterized according to relevant criteria, which commonly include detection performance, labour capacity, minimum survey intervals, environmental constraints, reporting delays, and follow-up criteria for screening methods. These templates can also be used to develop modules for other methods.

The quality of the parameterization depends on the empirical data available. For example, detection performance can be a static minimum detection limit or a function of leak size and environmental conditions. Labour capacity is defined by the number of unique agents (crews) that can be deployed, the duration of a workday (including daylight constraints), the time to complete each facility, and time spent between facilities, which combines transit, permitting, breaks, and complications. For reg-LDAR, minimum survey intervals reflect policy requirements designed to spread out surveys. For alt-LDAR, these intervals represent clearly-defined work practices within a proposed program. Reporting delays represent the time interval from detection to actionable information. Delays may be trivial for some methods, but functionality has been included for methods that are slow to process and deliver actionable information.

Environmental conditions can be incorporated in two ways. First, operational envelopes are defined, preventing deployment if conditions are unsuitable (a binary determination as defined by each method's work practice). Second, detection performance within an envelope can be configured as a function of the environment through a custom model relating environment to performance. As environmental conditions vary in space and time, so should the performance of LDAR methods. To account for local environmental context, LDAR-Sim uses gridded ERA-Interim reanalysis data. These data have global daily coverage with ~80 km spatial resolution from the European Centre for Medium-Range Weather Forecasts (Dee et al., 2011). At a minimum, LDAR-Sim requires temperature, wind speed, and precipitation as inputs, as these are known to impact the performance of OGIs,(Ravikumar and Brandt, 2017) which are typically used as the reference target in equivalence (AER, 2018; Government of Canada, 2018). However, the reanalysis contains a detailed description of the weather, including wind direction (for downwind plume intersection methods), snow depth and cloud cover (for remote sensing methods), and solar irradiance, which can be incorporated as needed.

Facilities in the LDAR program (henceforth 'sites') are provided to LDAR-Sim as geographic coordinates to establish local environmental context and topology. A single site could be used, a site list comprised of a company's assets, or all sites in a basin or jurisdiction. Custom survey frequencies can be assigned to each site for each method used in a program. These can be uniform for simple scenarios or based on facility type or risk factors such as production, age, or site-specific historical LDAR data. Survey frequencies can also be set according to accessibility. For example, if a site is not accessible by MGL, it can be designated an OGI site. The time required to survey each site can also be specified by method.

4.3.2 Initialization

To start the simulation, methods and site emissions are initialized. In LDAR-Sim the user can simultaneously deploy multiple methods, each with one or more independent crews. To achieve this structure, each method has a 'company' that manages daily work requirements by deploying crews. This construct also enables optimization of crew management. Each company in the program is initialized by building a deployment day matrix defining whether each site is available on each day of the simulation according to the environmental constraints set during program definition. This results in n_m matrices of dimensions n_s and t, where n_m is the number of methods in the program, n_s is the number of sites, and t is simulation timesteps in days. To register a deployment day, all environmental conditions at a site must satisfy the user-defined operational envelopes for that method.

Site emissions are initialized as fugitive and vented emissions. For each site, the number of initial leaks is randomly drawn from an empirical distribution of leak counts, and each leak is assigned an emission rate from an empirical leak-size distribution. Bias can arise if count and rate distributions are not acquired using the same detection method. For example, if counts are acquired from a study using Method 21 instruments and rates from a study using OGIs, aggregate emissions will be overestimated due to the lower sensitivity of OGIs. Ideally, both count and rate data should come from the same study and be representative of the region or assets of interest. If local data are limited or unavailable, inputs should be as representative as possible, and sensitivity should be tested.

Most LDAR programs target only fugitive emissions, but allowable vented emissions matter because methods that measure site-level emissions are often unable to distinguish between the two. Previous studies and inventories have commonly attributed > 50% of site-level emissions to vents (Allen et al., 2013; Johnson et al., 2017), so it is important to consider the potentially confounding effects of these sources. To date, 'bottom-up' component-level measurements have been difficult to reconcile with contemporaneous 'top-down' site-level measurements (Brandt et al., 2014). Possible explanations for the large discrepancy include (1) inability to measure certain sources due to access restrictions (e.g. tank vents) or instrumentation limitations (e.g. compressor exhaust lacks quality measurement methods), and (2) increased chance that top-down methods capture episodic events such as blowdowns and liquid unloadings (Johnson et al., 2017; Zavala-Araiza et al., 2015). Rather than attempt to estimate site-level vented emissions (V_i) from component-level measurements, LDAR-Sim takes an empirical set of site-level measurements and bootstraps the count (C), leak (L), and site-total (T) empirical distributions to estimate site-level vented emissions:

$$V_i = T_i - \sum_{j=1}^{C_i} L_j.$$

This procedure generates a distribution of i (typically 1000) site-level vented emission estimates. Whereas leaks are assigned to sites and remain at the same site, vents are assigned stochastically during each survey to simulate the episodic and dynamic nature of vented emissions (Johnson et al., 2019; Vaughn et al., 2018).

4.3.3 Updating

After initialization, LDAR-Sim evaluates the program at daily timesteps. Tracking the passage of time is of fundamental importance in LDAR-Sim and relies on empirical inputs. Each daily timestep is broken down to the minute scale. The primary time control mechanisms are: (i) the time it takes each method to survey each facility; (ii) the time it takes for agents to travel between facilities; (iii) a user-defined maximum number of work hours per day; and (iv) optional daylight requirements that calculate daylight hours for a given latitude and day of the year. Several user-defined delays can also be introduced, including reporting delays (how long it takes for survey information to become actionable), repair delays (how long it takes for a leak to be repaired once it has been reported), and minimum survey intervals. Each day can be split into three parts: adding new leaks, finding leaks, and repairing leaks. New leaks (L_n) are generated using a site-level empirical leak production rate (LPR; leaks · site^{-1} · day^{-1}) that is independent of the number of leaks already present on site:

$$L_n \sim Binomial(1, LPR)$$

The LPR is an empirical representation of all conditions that lead to the occurrence of leaks, including facility age, management practices, and predictive maintenance. In theory, each site can have $t + L_i$ leaks, where L_i is the initialized leak count. This is unrealistic and is made impossible in FEAST, which limits the number of leaks at a site to a finite component count. However, leak removal processes described in the following section prevent unrealistic accumulation of leaks, and we find that specifying an upper limit is unnecessary. In LDAR-Sim only one new leak can be added to a site per day. However, for a typical LPR of 0.006, there is only a 3.6 x 10^{-5} chance that two new leaks will arise at a site on the same day (approximately once every 65 years); we do not consider this an important concern.

In LDAR-Sim, we distinguish between close-range methods that are able to confirm and diagnose leaks at the component scale, and screening methods that measure at larger spatial scales. Leaks must be detected at the component scale before they can be tagged for repair and enter the repair queue. This can happen in two ways: by operators during routine visits to their sites or with close-range LDAR (e.g. OGI). Methods measuring at the equipment or facility scale (e.g. MGLs, aircraft, satellites, and most drones) must flag sites for follow-up by a close-range method as they are generally unable to discern individual sources or distinguish between vented and fugitive emissions. Operator detection represents periodic inspection, maintenance, and repair outside of an LDAR program and is governed by a null repair rate (NRR). An overview is presented in Section 7.1 with further discussion in Sections 4.3.4 and 4.5 below.

When an LDAR program is configured, each company in the program is instructed to work each day. Companies then deploy each of their crews and leak detection occurs in accordance with the method's work practice, user-defined performance metrics, and program requirements. When a crew is deployed, workday duration is first calculated as the shorter of either (1) the number of daylight hours, or (2) the user-defined maximum workday duration for that method. Daylight hours are calculated using the ephem package in python, which returns civil twilight sunrise and sunset times for a given date, latitude, and longitude (Rhodes, 2011). The crew then 'chooses' and 'visits' sites until the workday ends.

Crews attempt to visit sites that have gone the longest without LDAR. A series of checks ensures that (1) the minimum interval has passed between surveys, (2) the site is not yet in compliance for the year (to avoid conducting too many surveys), and (3) environmental conditions fall within the operational envelope. If any of these conditions are not met, the crew considers another site. If no suitable sites are identified, the workday ends. If a site is identified, the crew visits it and each leak is subject to the crew's detection module. Detected leaks are tagged for repair. Once all leaks have been attempted, the time of day for the crew advances by the sum of: (1) the site-specific survey time for the crew's method, and (2) an estimate of the time required to reach the next site. Screening method templates are similarly configured, but flag entire sites for follow-up by a close-range method. When a site is surveyed, the method's detection module is applied to the cumulative site emissions (Q_{site}), which consists of all leaks and a vented emissions estimate (V) drawn from the empirical distribution of vents:

$$Q_{site} = V + \sum_{l=1}^{n_{leaks}} Q_l$$

If a site is detected it can be flagged, but only if Q_{site} exceeds an optional user-defined follow-up threshold. The purpose of the threshold is to prevent highly sensitive screening methods from sending follow-up crews to each site they screen. A 'follow-up ratio' may also be specified, which flags sites based on relative emissions. For example, a follow-up ratio of 0.2 will flag the top 20% of surveyed sites.

Leaks are repaired following optional reporting and repair delays. All methods have a reporting delay separating data collection from delivery of actionable information. For screening, two reporting delays exist in the flag-tag-repair workflow as a liaising follow-up crew is required. Once tags are reported, a repair delay must pass. The repair delay can be a single value or an empirical distribution. The true distribution is unknown but may be skewed or bimodal as some leaks can be repaired immediately (e.g. tightening a valve) while others require an equipment shutdown.

4.3.4 Finalization

When all simulation days are complete, LDAR-Sim outputs a series of data files, maps, and figures. Data are output by site, leak, and day to enable detailed and transparent evaluation of results (Table 7.1). The proportion of sites available is a global measure of suitability reflecting the daily proportion of all sites that fall within each method's environmental constraints. Although cost is not presently treated thoroughly in LDAR-Sim, users are encouraged to build their own costing modules with real data. Our placeholder cost variable is a crude estimate of daily company costs and is charged for each crew deployed.

When multiple methods are deployed, it is important to track different kinds of detection redundancy that may arise. Redundant tags are tagged leaks that are detected a second time by an operator or close-range method. These tags are not counted, as the leak is already waiting repair. Three types of redundant flags exist and are described in Figure 7.1. Tracking redundancy is important to ensure that multiple companies or crews are not taking credit for emissions reductions resulting from the repair of the same leak.

In principle, two LDAR programs could be compared using various metrics. Examples include average emission rates, active and repaired leak counts, missed leaks, average duration between leak origin and repair, and so on. Here, our focus is on emissions reduction equivalence, the metric commonly prescribed by regulators. Various approaches can be used to estimate emissions reductions. In FEAST, emissions under an LDAR program are subtracted from a null scenario – 'business as usual' in the absence of LDAR (Ravikumar and Brandt, 2017). The null scenario balances LPR and NRR in the absence of LDAR to achieve steady-state emissions centered on the initialized state of the leak field.

We find the notion of a null scenario problematic for several reasons. First, the 'pre-LDAR' baseline is elusive, and uncertainty in its estimate propagates to mitigation. Today, many companies have some kind of leak management program already in place. In Canada, a shift towards more formalized LDAR occurred in 2007 when best management practices were published by an industry group.(CAPP, 2007) Second, fugitive emissions are unlikely to be in equilibrium. Management practices, new technology, and pressure on producers to reduce emissions may result in a downward trend in fugitive emissions. While both possibilities are speculative, the burden of proof should be on equilibrium if this assumption is required. Third, the null baseline is maintained by LPR and NRR, two poorly understood parameters. Both LPR and NRR have significant challenges:

- LPR has never been estimated independently of NRR; in deriving LPR, it is assumed that NRR and LDAR are inconsequential. This is inconsistent with assumptions made during simulation, where NRR is assigned a value approximately equal to LPR. If estimating an LPR that already encodes null repair, NRR should not need to be applied during simulation.
- LPR is most easily estimated from data acquired during LDAR surveys, compromising the 'pre-LDAR' steady state assumption.

- While attempts have been made to estimate LPR, NRR has neither been measured nor empirically inferred.
- In FEAST, NRR is a static value defined using LPR. The same NRR is used for all leaks, regardless of emissions rate. However, large leaks may be more easily detected by human senses, and if discovered, are of greater safety and financial concern. If removal rate is a function of leak size, production of larger leaks must increase to maintain the relative proportion of large leaks.
- NRR likely depends on company culture and individual operator practice. In FEAST, a higher LPR results in a higher NRR. However, if company culture matters, LPR and NRR may be inversely related.

Given these assumptions, it is conceivable that the best guesses for LPR and NRR are wrong, perhaps by a large margin. We therefore avoid the concept of a null scenario and compare programs directly based on emissions instead of mitigation. This can be formalized: Let equivalence between two programs (P_1 and P_2) be defined as equal mitigation ($M_1 = M_2$). Assume that $M = Null_E - P_E$, or emissions from the null scenario minus program emissions. If the null scenario is the same for both programs, equivalence can be defined as $Null_E - P_{E1} = Null_E - P_{E2}$, where $Null_E$ cancels out.

4.4 Case studies

To demonstrate LDAR-Sim, we present two case studies of LDAR programs deployed in Alberta, Canada. In the first case study, two alt-LDAR programs are compared to a hypothetical reg-LDAR program. The first alternative program (P_1) uses quarterly MGL screening combined with follow-up OGI at 80% of facilities emitting more than 100 $g \cdot h^{-1}$ (Table 4.1). The second alternative program (P_1) requires aircraft screening 6 times per year with follow-up OGI at 50% of the facilities emitting more than 5000 $g \cdot h^{-1}$ (Table 4.1). The reference reg-LDAR program (P_{Ref}) requires triannual OGI at all facilities. In the second case study, PRef is compared to three additional OGI-based programs (P_W, P_L, and $P_{W,L}$) to explore how reg-LDAR program performance can vary with environmental and labour constraints. The three new OGI programs differ from PRef in either (i) temperature, wind, and precipitation deployment thresholds (P_W and $P_{W,L}$), or (ii) the number of deployment crews available to do work (P_L and $P_{W,L}$). P_W, P_L, and $P_{W,L}$ are otherwise identical to P_{Ref}. All case study programs presented in this paper are hypothetical and intended only for demonstration.

All case study programs are implemented over a 7-year period (2011-2018; using historical weather data) alongside the OGI-based regulatory standard required by Environment and Climate Change Canada (ECCC) (Government of Canada, 2018). Briefly, ECCC requires triannual surveys on most upstream production facilities, with the exception of single wellheads with no associated equipment (i.e. connected directly to a gathering pipeline). Most leaks must be repaired within 30 days, unless a major shutdown is required. We demonstrate these programs for a producer with 1169 Alberta facilities requiring LDAR. To maintain anonymity of the producer, random noise $\mathcal{N}(0, 0.1)$ was added to the x and y coordinates of each site and no identifying information is reported. This has negligible effect on the conditions polled from weather data. The sites are broadly representative of Alberta's diverse geography and producing regions, spanning approximately 9.6° of longitude and 8.7° of latitude. For each program we conduct 25 simulations over 2557 daily timesteps and assume a standard repair delay of 14 days for all leaks. Input parameters are summarized in Table 4.1 and explained in Section 7.2.

Case study results are presented in Figure 4.2 and Table 7.2. In general, emissions in Figure 4.2 increase according to the LPR and decrease when LDAR is conducted. Each program configuration leads to distinct emissions trends (Figure 4.2). Only P_1 achieves equivalence to the reference case. Although less sensitive (the OGI median detection limit is ~10 $g{\cdot}h^{-1}$), P_1 measures cumulative site-level emissions, including vents, and sends follow-up crews to 80% of detected sites. In this configuration, P_2 is not equivalent to P_{ref} as it is only visiting sites emitting above 5000 $g{\cdot}h^{-1}$. Small leaks are still eventually resolved, as follow-up crews conduct comprehensive surveys when tasked to a high-emitting site. Despite large differences in emissions (and therefore mitigation), each of the programs results in roughly the same number of tagged leaks (Table 7.2). This suggests that screening methods may be less efficient than close-range methods, because (1) the presence of vented emissions can influence follow-up decisions, and/or (2) higher detection limits require more leaks to build up before a follow-up is triggered.

Periodicity in emissions is driven by survey frequency, labour availability, and deployment constraints. The slope of emission reductions relates to the size of the workforce relative to the number of facilities to be surveyed. In FEAST, this slope is vertical, creating a 'saw tooth' emissions profile not observed in LDAR-Sim because agents typically require several weeks or months to complete each round of surveys. The absence of periodicity in P_L and P_1 suggests that agents are working continuously. The triannual fluctuations in P_{ref} suggest that crews survey all required sites before the next round of surveys is required. A balance must be found between having idle crews and a risk of falling out of compliance. Strong periodicity can be indicative of an oversupply of labour or of operational envelopes restricting surveys to seasonal clusters. The weather-restricted programs (P_W and $P_{W,L}$) exhibit strong seasonal (wintertime) periodicity resulting in ~50% higher fugitive emissions (Figure 4.2C, Table 7.2). To better understand how weather affects deployment days, LDAR-Sim generates maps to explore the spatial distribution of suitable conditions for a given set of environmental constraints (Figure 7.2).

The variable performances of the OGI-based programs in Figure 4.2C illustrate the importance of defining the reference program for determining equivalence. Most jurisdictions require demonstration that alt-LDAR is equivalent to reg-LDAR but are not explicit about the quantity of emissions permitted or mitigated. When P_2 is compared to P_{ref} it does not achieve equivalence. However, P_2 is closer to $P_{W,L}$ and could be deemed equivalent. To achieve defensible definitions for emissions reductions, regulators should be explicit in delineating permissible contexts or expected mitigation for equivalence demonstration.

Table 4.1 Parameters used for the case study. These programs are hypothetical and based on fabricated detection methods and survey requirements for the sake of demonstration.

	P_{Ref}	P_W	P_L	$P_{W,L}$	P_1*	P_2*	†
Method	OGI	OGI	OGI	OGI	MGL	Aircraft	OGI
Number of crews	3	3	1	1	1	1	2
Minimum temp (°C)	-30	0	-30	0	-30	-30	-30
Maximum wind speed ($m \cdot s^{-1}$)	20	5	20	5	20	20	20
Maximum precipitation (m)‡	0.01	0.001	0.01	0.001	0.01	0.01	0.01
Minimum repeat interval (days)	120	120	120	120	50	50	-
Maximum workday (hours)	10	10	10	10	10	10	10
Reporting delay (days)	2	2	2	2	2	2	2
Detection limit ($g \cdot hour^{-1}$)	a	a	a	a	100	5000	a
Time offsite ($minutes \cdot site^{-1}$)	b	b	b	b	b	10	b
Required surveys per site	3	3	3	3	4	6	-
Time per site (minutes)	120	120	120	120	30	5	120
Follow-up threshold ($g \cdot hour^{-1}$)	-	-	-	-	100	5000	-
Follow-up ratio	-	-	-	-	0.8	0.5	-

* Screening programs include two OGI follow-up crews

† This column gives relevant follow-up parameters for screening methods.

‡ Accumulated between 12:00 and 18:00.

[a] Configured using Ravikumar et al (2018). See text for details.

[b] Drawn from an empirical distribution. See text for details.

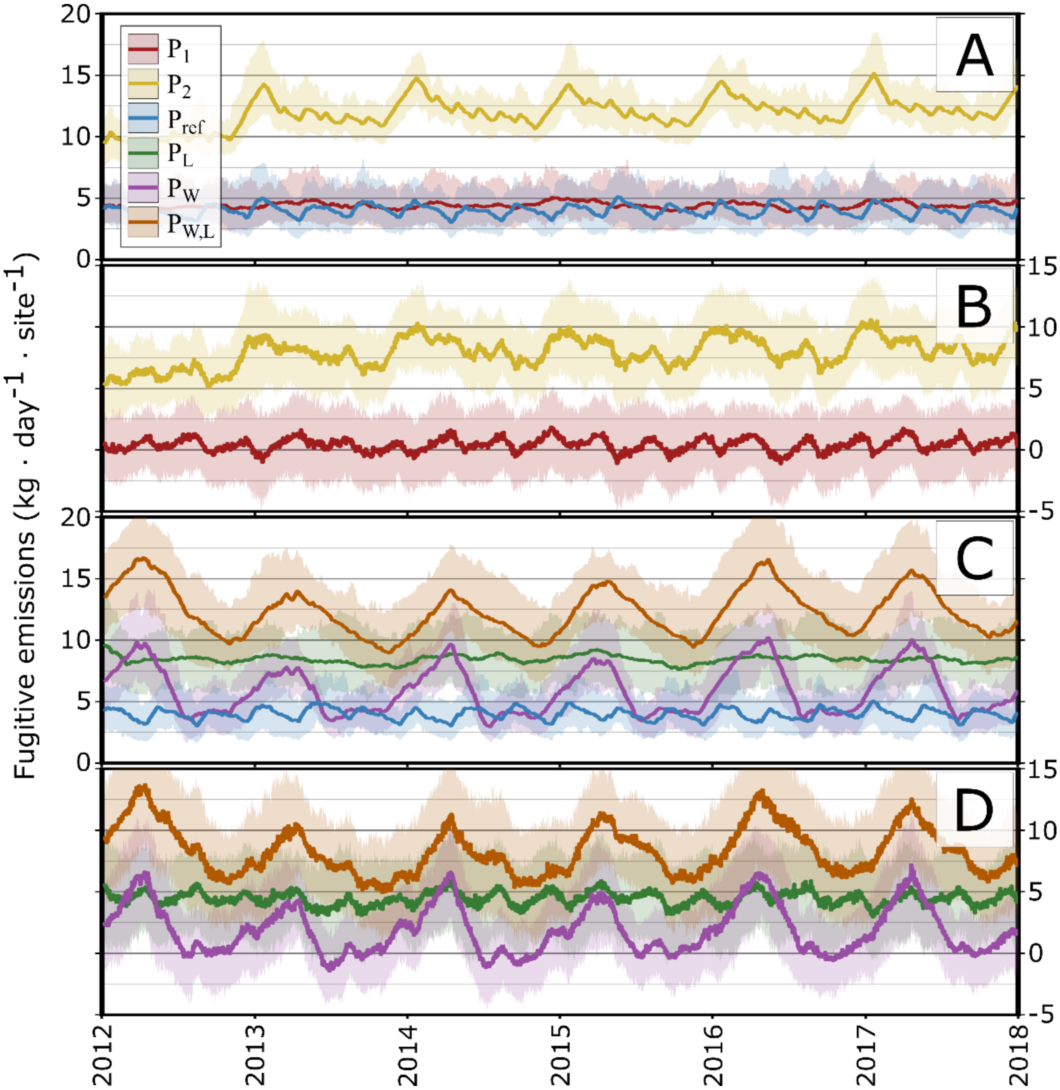

Figure 4.2 Fugitive emissions observed under each program. Lines show daily emissions averaged across 25 simulations (± 2σ). A and C show average site-level emissions, while B and D show candidate programs differenced against the regulatory program ($P_X - P_{ref}$). In B and D, equivalence occurs at y = 0.

4.5 Model confidence

Building confidence in LDAR models is challenging because (1) many of the parameters used to characterize leak fields, detection methods, and repair remain poorly constrained, (2) empirical inputs, and therefore results, vary in time and space, and (3) validation capacity is limited by the high cost of LDAR programs, the long periods over which they can be implemented, and the challenge of accurate emissions measurement. Although the value of these tools is heuristic, model confidence can be established by comparing similar models, conducting SA, and comparing against real LDAR programs.

The only model that can be used for comparison is FEAST, which has similar leak field properties but differs in most other aspects. Results from FEAST simulations of OGI-based LDAR programs are presented in two studies.(Kemp et al., 2016; Ravikumar and Brandt, 2017) Both use the same LPR as our case study, but all three studies use different empirical leak distributions. In the first, fugitive emissions fluctuate between ~1-10 $kg \cdot day^{-1} \cdot site^{-1}$ for a survey interval of 100 days.(Kemp et al., 2016) In the second, quarterly and annual surveys result in mean emissions rates of ~13 and 26 $kg \cdot day^{-1} \cdot site^{-1}$, respectively.(Ravikumar and Brandt, 2017) Results from LDAR-Sim's OGI demonstration fall within these two studies (~3-13 $kg \cdot day^{-1} \cdot site^{-1}$; Figure 4.2), although neither model has been validated against field measurements. To explore how different parameterizations could impact equivalence determinations, we compare an OGI-based program to a generic MGL screening program in an extensive SA (see Section 7.3 for methods). We compare the two LDAR programs using both differences and ratios of their outputs to explore which method provides more robust equivalence metrics.

Results show that leak production rate is the most sensitive input parameter (Figure 4.3), similar to past results (Kemp et al., 2016). As LPR increases, the absolute difference in emissions between programs increases (Figure 4.4A), with the OGI program outperforming the screening program. However, relative emissions between programs appear independent of LPR (S = -0.3 for ratio and 1.7 for absolute difference; Figure 4.3), suggesting that detection capacity of different programs may scale with total emissions, despite falling behind in absolute terms. Given that LPR remains an elusive parameter, we recommend comparing programs on a relative rather than absolute basis. If absolute differences in emissions are used, establishing equivalence may be easier for low-LPR companies or regions. Future research of repeat field surveys should feature robust record keeping of repairs and leak history (new, re-occurred, and not repaired) to improve LPR estimates.

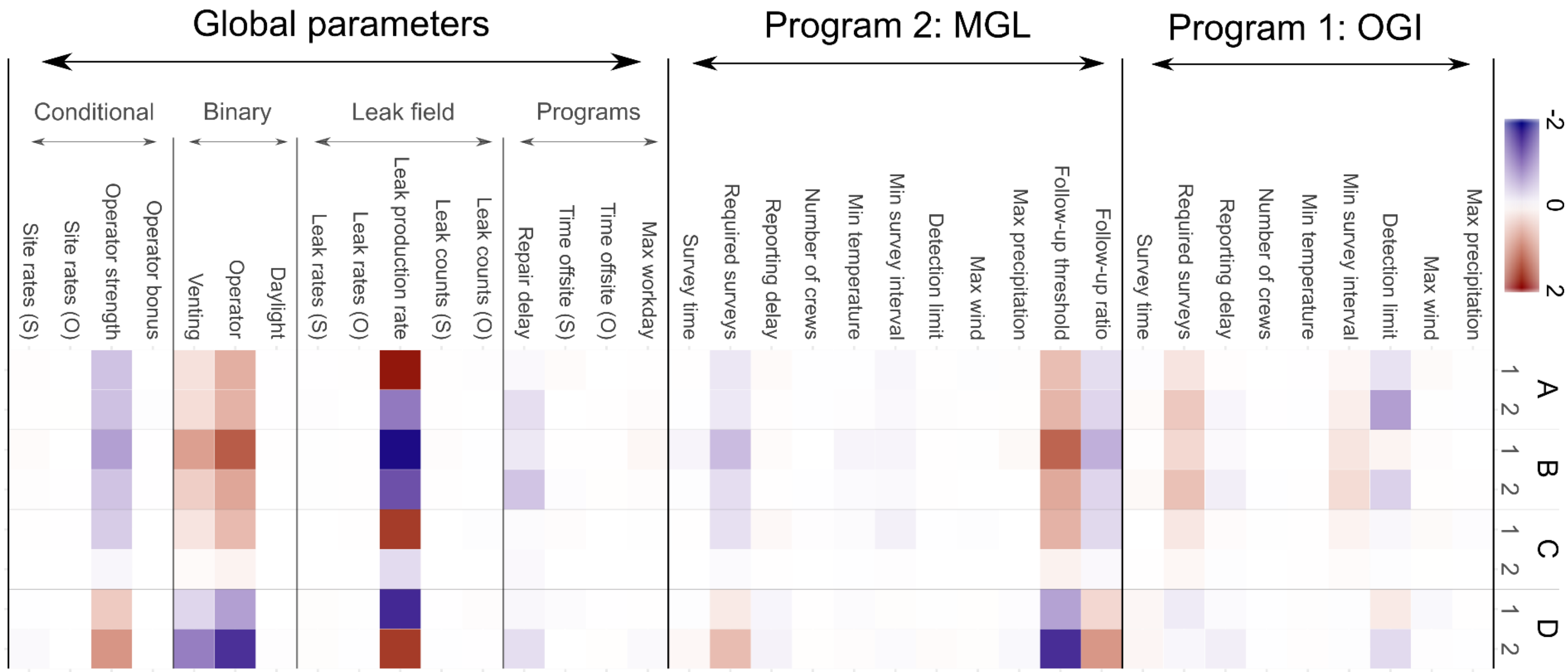

Figure 4.3 Sensitivity analysis results. Darker colours indicate more sensitive inputs; see Section 7.3 for derivation of sensitivity index $\boldsymbol{S}$. Outputs are: (A) median number of active leaks, (B) median number of days leaks remain active before discovered, (C) mean daily site emissions, (D) cumulative number of repaired leaks. For each output, column 1 calculates $\boldsymbol{S}$ using absolute differences between programs ($\boldsymbol{MGL - OGI}$), while column 2 uses a ratio ($\boldsymbol{MGL/OGI}$). Values of $\boldsymbol{S}$ near zero suggest that the input parameter has little influence over the equivalence metric, though don't necessarily indicate that the program is unaffected by the input.

Follow-up work practices and the presence of venting impact the performance of screening programs (Figure 4.3). Screening methods that measure site-level emissions are unable to distinguish between vented and fugitive sources. As vented emissions increase, follow-up decisions may preferentially target high-venting sites at the expense of fugitive emissions. The impact of venting on relative performance is higher when LPR is low (Figure 4.4B), likely because fugitive emissions are lower and more likely to be overwhelmed by vented emissions, which are independent of LPR. The SA suggests that increasing follow-up ratios and decreasing thresholds can improve screening performance. For alt-LDAR programs to be appealing to producers, trade-offs between follow-up rules and deployment costs should be considered.

Results are sensitive to assumptions surrounding operator activity (Figure 4.3). As more leaks are repaired by operators, fewer remain for LDAR programs. Higher operator activity therefore results in smaller differences between programs, making demonstration of equivalence easier. The operator effect tends to increase with LPR, but mostly when comparing absolute differences in outputs. Given the evidence available, we advocate for disabling NRR in most situations, because (1) most empirical LPR estimates implicitly encode NRR, making additional NRR activity redundant; (2) no empirical NRR data exist and AVO detection metrics have not been established; (3) the assumption that system-wide fugitive emissions are in equilibrium lacks evidence; and (4) incorporating NRR obscures program comparison, increasing the threat of erroneously declaring two programs equivalent. Although LPR and NRR are nearly impossible to measure separately, they are relatively easy to measure together. Rather than attempt to estimate both LPR and NRR independently, we recommend modelers (1) reject the requirement of a steady-state and a null scenario and (2) work only with the empirical LPR that already accounts for NRR.

Several input variables are of seemingly low influence (Figure 3). For example, LDAR-Sim outputs do not appear to be sensitive to the addition or removal of super-emitters in any of the empirical input distributions. Others include operator bonus (p_max), detection limits for MGLs, and leak-size distributions. However, these inputs may still impact the performance of individual programs (e.g., Figure S5). For example, increasing the number of super-emitters can increase program emissions, but if the increase is approximately proportional for both programs, their relative performance remains similar and S will be near zero. Additional variables and functionality not considered in our case study or SA may be important, depending on program configuration. Targeted SAs are therefore necessary each time LDAR-Sim is used, as input variables may become sensitive when tested against a narrower range of other inputs.

At present, collective understanding of how to configure LDAR programs is nascent. Many of the inputs required to define LDAR programs are poorly constrained. Detection modules presented here are based on hypothetical performances and are not tied to specific work practices. To be used effectively, custom modules should be developed for each candidate method. These modules should reflect each method's unique set of performance metrics developed through controlled testing. Field data will improve understanding of these variables and identify others that may warrant inclusion. By means of open-source development, LDAR-Sim can evolve alongside this knowledge base, and may grow to be used in different ways by regulators, producers, consultants, solution providers, technology developers, investors, and non-profits.

Regulators can play an important role in this transition and enabling the success of tools like LDAR-Sim. Not all jurisdictions have introduced policies to mandate monitoring and reduction of fugitive methane emissions. Of those that require LDAR, few permit alt-LDAR. However, interest is growing rapidly and more alt-LDAR policies are expected soon. As new policies emerge, regulators should work together to establish clear requirements and standardized protocols for demonstrating equivalence, especially where modeling is involved. Regulators should also work to clearly and quantitatively define equivalence based on emissions reduction targets and expected performance of approved methods. If authorizing the use of LDAR-Sim, regulators should also be intimately familiar with the ways it might be used by applicants. Ultimately, LDAR-Sim is one tool that can be used as part of a broader effort to identify emerging LDAR methods and should be supported with empirical data and expert advice.

LDAR-Sim is a modular, open-source modeling framework that can be further developed as new functionality is required or new measurement methods are developed. Future work will identify alt-LDAR programs that improve upon reg-LDAR, track fugitive and design emission with high spatial and temporal resolution, develop methods to optimize cost to mitigation ratios, validate model predictions through field trials, and improve integration of empirical inputs. More flexibility in spatial resolution should be built into LDAR-Sim, as some methods monitor at the equipment scale, providing more detailed information than facility-scale screening. Other possible improvements include modeling predictive maintenance efforts, quantification accuracy of screening methods, additional environmental variables to define operational envelopes of new methods, and a more sophisticated reproduction of repairs.

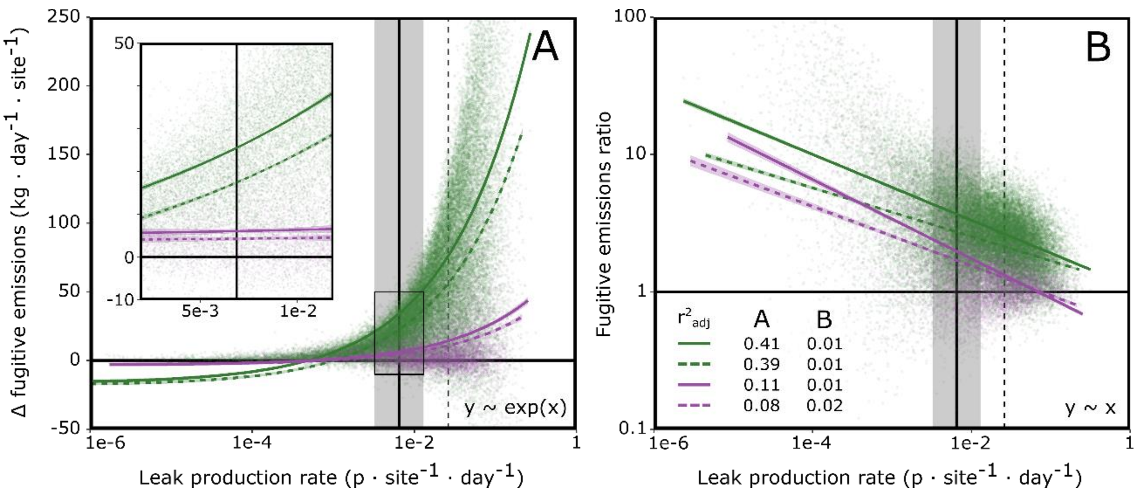

Figure 4.4 Differences (A; MGL - OGI) and ratios (B; MGL/OGI) of fugitive emissions as a function of LPR. Each dot compares a close-range and a screening program that share global parameters. Purple and green lines show operator presence and absence, respectively (i.e. NRR turned on or off). Dashed lines show presence and solid lines show absence of vented emissions. The dashed vertical line is the mean LPR from the SA sampling distribution (0.026), established using Alberta data. The solid vertical line is the LPR value used in FEAST and the shaded region is the LPR SA range evaluated in Kemp et al. (2016)

Chapter 5

Will new mobile screening technologies enable cost-effective methane leak detection?

5.1 Abstract

New methane-sensing technologies such as drones, satellites, and aircraft have emerged to help identify and reduce fugitive emissions from the oil and gas sector as part of leak detection and repair (LDAR) programs. Many of these new technologies use multi-visit LDAR (MVL), which consists of four steps: screen all facilities, triage, follow-up, and repair. First, rapid screening surveys are used to detect and/or quantify facility-scale emissions. Next, a triaging procedure selects a subset of facilities for follow-up. Finally, follow-up inspectors are dispatched to identify, diagnose, and ultimately repair individual leaks. Regulators require that MVL achieves equivalent mitigation of fugitive methane emissions as compared to a regulatory (e.g., OGI only) program. Producers are more likely to adopt MVL if it is also more cost-effective than a regulatory program. We use the Leak Detection and Repair Simulator (LDAR-Sim) to explore whether MVL can achieve equivalence while being more cost-effective than incumbent methods. Results show that facility-scale MVL programs can achieve equivalence, but at high cost. Under a best-case scenario with (i) no screening quantification error and (ii) no design emissions (allowable venting or combustion) at facilities, we find that MVL must cost less than 100 USD per facility to achieve 30% cost reductions relative to regulatory OGI. In more realistic scenarios, design emissions and screening quantification uncertainty lead to triaging errors that force excessive and poorly deployed close-range follow-up. In these scenarios, the cost of MVL exceeds regulatory OGI. Our work suggests that decision-making procedures for directing follow-up are crucial for cost-effective MVL. Design emissions are currently the biggest challenge facing MVL, suggesting a regulatory shift towards monitoring all emissions, not just fugitives, could improve the viability of screening.

5.2 Introduction

The oil and gas (O&G) industry is among the largest sources of anthropogenic methane, a potent greenhouse gas. To reduce fugitive (i.e., unintentional) methane emissions from O&G, regulators are increasingly mandating the use of leak detection and repair (LDAR) programs (AER, 2018; ECCC, 2018; EPA, 2016). Most LDAR programs rely on handheld organic vapour analyzers

(OVA) or optical gas imaging (OGI) cameras, which are sensitive, familiar, and recommended by regulators (EPA, 2016). However, rapid innovation is underway; diverse new LDAR technologies and methods are emerging (Fox et al., 2019a; Ravikumar et al., 2019). Regulators in several jurisdictions have opened paths to approval for these alternatives. For example, in Alberta, Canada, O&G producers can propose any alternative LDAR program, but to be approved it must reduce the same quantity of fugitive methane as a prescribed regulatory program (AER, 2018; Fox et al., 2019b). Similar requirements for emissions reduction equivalence exist in other jurisdictions that enable alternative LDAR (CDPHE, 2018; ECCC, 2018). Most regulatory programs consist of exhaustive inspections at all facilities using OVA or OGI. Fundamentally, alternative LDAR programs must have: (i) equivalent emissions reductions to the regulatory program (to be approved by regulators), and (ii) lower cost than the regulatory program (to be of interest to operators).

Emerging LDAR technologies differ markedly in measurement scale, deployment mode, and information product, leading to a range of niche use cases (Fox et al., 2019a). Many new technologies consist of methane sensors deployed on mobile platforms: vehicles (Robertson et al., 2017; Caulton et al., 2018), drones (Golston et al., 2018; Barchyn et al., 2019), aircraft (Englander et al., 2018; Schwietzke et al., 2018), and satellites (Jacob et al., 2016; Varon et al., 2018). Data gathered by LDAR technologies can be used to attribute emissions to various scales, including regions, facilities, equipment, or components. Ultimately, component-scale detection is required for diagnosis and repair, but most emerging technologies measure at equipment or facility scales. A new strategy called ‘screening’ has therefore emerged, which uses rapid mobile surveys to identify a subset of the highest-emitting facilities in greatest need of LDAR (Fox et al., 2019a). The motivation behind screening is to reduce the cost of LDAR by only targeting the highest emitters. Studies in Canada and the US have repeatedly shown that most methane leaks are small, with 5% of sources accounting for approximately 50% of emissions (Brandt et al., 2016). Proponents of screening argue that if the largest can be found quickly, many of the smaller and relatively inconsequential leaks can be overlooked.

Screening must be combined with close-range follow-up inspections to diagnose and repair individual leaks. We define multi-visit LDAR (MVL) programs as consisting of four steps: screening, triaging, follow-up, and repair. First, rapid screening assesses all facilities, detecting and often quantifying emissions. Second, triaging is a decision-making step that uses the results from screening to determine which facilities should receive follow-up inspections. Most MVL programs consist of a single round of screening, but screening and triage can be repeated if multiple methods of increasing precision are used (e.g., satellite then aircraft). Third, facilities that were flagged during triaging are inspected at close-range to confirm whether a leak exists and identify repair requirements. Fourth, if a leak exists and cannot be immediately repaired, an additional visit may be required. In general, the goal of MVL is to reduce time-consuming and expensive OGI surveys at every facility by combining cheaper screening with targeted follow-up inspection at a subset of screened facilities. Screening frequency in MVL is generally higher than inspection frequencies in the regulatory program because (1) most screening methods are less sensitive than close-range methods, and (2) follow-up occurs at only a subset of facilities, while regulatory programs include all facilities.

Certain conditions are favourable for MVL. First, screening costs per facility should be lower than close-range inspections because (i) screening is applied more frequently, and (ii) after screening, follow-up surveys may still be required. Second, leak-size distributions with heavier tails (more small leaks, fewer large leaks) favour MVL because the benefit of finding large leaks increases. Third, deployment conditions and technology performance should enable accurate triaging. When ranking facilities by emission rate, quantification error and the presence of confounding sources may affect triaging. For example, design emissions (i.e., allowable venting and fuel slip during combustion) at upstream facilities are not classified as leaks and do not fall under LDAR regulations yet are measured by facility-scale screening technologies. Quantification error from screening methods is large (Ravikumar et al., 2019). More accurate quantification is often possible with longer surveys – but screening must be rapid to be economically viable. When triaging decisions are inaccurate, the ranked list of facility-level emissions is inaccurate, and lower-emissions facilities can get visited at the expense of higher-emissions facilities. In the worst-case scenario, follow-up inspectors could be sent to facilities with no leaks, while facilities with very large leaks are overlooked.

Here we examine whether MVL can be both equivalent and cost-effective. Using an agent-based modeling framework called LDAR-Sim, we generate a broad range of equivalence scenarios, spanning different screening survey frequencies, follow-up requirements, quantification errors, empirical leak-size distributions, and design emissions intensities. These equivalence scenarios are then used to explore the cost-effectiveness of MVL under different conditions. We show that equivalent and cost-effective MVL programs require extremely low-cost screening and provide target metrics to help identify successful technologies and programs.

5.3 Methods

Equivalence and cost-effectiveness for MVL are evaluated using the LDAR Simulator (LDAR-Sim; Fox et al., in review). LDAR-Sim is an agent-based numerical modeling framework that enables precise definition of LDAR programs. In a virtual asset field, LDAR workers search for leaks at O&G facilities. LDAR workers apply technology modules that can detect leaks or quantify facility-scale emissions. The deployment of LDAR workers is managed under a program definition, which defines the number of workers and types of technology they are using. New leaks and design emissions appear stochastically in the asset field, drawn from empirical distributions. The model proceeds forward through time, tracking the performance of different LDAR programs. Empirical inputs describing specific deployment regions, target facilities, monitoring technologies, work practices, and regulations are required to operate the model. Outputs describe anticipated emissions mitigation and a cost analysis of the proposed program. To demonstrate equivalence, simulations can compare regulatory and alternative LDAR programs. Please see Fox et al. (in review) for a comprehensive description of LDAR-Sim.

In LDAR-Sim, inspection agents can measure emissions at the component scale, simulating handheld instruments like OGI, or at the facility scale, simulating screening methods. In this study we assume that facility-scale screening methods are unable to discern source type, so all fugitive and design emissions are aggregated into a single measurement. Some screening methods can measure emissions at intermediate scales, localizing sources to pieces of equipment (e.g., liquid tanks) or equipment groups but not individual components (Ravikumar et al., 2019). However, modeling at intermediate scales is complex and we leave these analyses for future work.

In LDAR-Sim, a MVL workflow consists of the following four steps: (i) Screening agents estimate emission rates of each facility in the LDAR program; (ii) A triaging procedure ranks all screened facilities by emission rate and flags those that meet a follow-up threshold; (iii) Follow-up OGI agents visit flagged facilities to identify individual leaks and tag them for repair; (iv) Leaks are repaired, lowering emissions. We use the following terminology: entire facilities are ‘flagged’ during screening and individual leaks are ‘tagged’ for repair. A flagged facility could have many small leaks, several moderate leaks, or one super-emitter (our simulated screening methods cannot differentiate these sources). A facility with no leaks may even be flagged for follow-up if design emissions are high, or if mistakes are made during ranking due to quantification error. However, follow-up surveys at these mis-ranked flagged facilities often still result in tags of smaller leaks.

The number and timing of follow-up inspections is important for determining fugitive emissions in LDAR programs. In MVL, emissions reductions are positively correlated with screening survey frequency and follow-up. Additional screening surveys will find newly generated leaks faster, reducing the time that a large leak is left emitting. More follow-up surveys will find more leaks. For example, an LDAR program with 500 facilities may have triannual screening and dispatch follow-up inspectors to the top 10% of emitters. Each year, 1500 screening surveys resulting in 150 flags and follow-up inspections will be conducted. To double the number of follow-up inspections to 300, three options are available: (i) double the number of screening surveys to 3000, (ii) double the percentage of facilities receiving follow-up to 20%, or (iii) some combination of (i) and (ii).

To better examine the economics of LDAR we vary screening frequencies and follow-up to isolate specific scenarios that achieve emissions reduction equivalence with regulatory programs. These 'equivalence scenarios' are used as the base of cost comparisons. Equivalence scenarios are important as regulators often approve alternative LDAR programs based on achieving emissions reduction equivalence with regulatory LDAR programs. Alternative programs will not be approved if they do not achieve equivalence, but those responsible for conducting LDAR (e.g., O&G producers) want to minimize costs and are unlikely to reduce emissions beyond what is required. Equivalence scenarios balance these opposing requirements.

This study consists of identifying equivalence scenarios and evaluating their cost-effectiveness. First, we run LDAR-Sim to estimate emission under regulatory annual and triannual OGI surveys. We then run ensembles of MVL simulations to identify all equivalence scenarios (Figure 5.1). Additional modeling parameters and assumptions affect results. We develop equivalence scenarios for two leak size distributions, two follow-up threshold definitions, with and without design emissions, and with and without quantification error. In part two, we evaluate the cost-effectiveness of equivalence scenarios for generic aircraft and truck-based screening. All analysis code and empirical inputs are available in the SI.

5.3.1 Triaging procedures

Triaging combines ranking facilities by emission rate and deciding which facilities to flag. Once facilities have been ranked by emission rate, an explicit rule must be used to distinguish between (i) high-emitting facilities that are flagged for follow-up inspections that tag leaking components and queue repairs and (ii) low-emitting facilities that will not receive close-range inspection and repair. Here, facilities are flagged if their estimated emission rate during screening is above a threshold emission rate. We introduce two types of follow-up threshold: static and dynamic.

A static follow-up threshold is a constant value derived from a baseline emissions distribution; it is the emission rate that corresponds with a desired target proportion of highest emitting facilities (static target). Table 4.1 shows the static targets used in this study, the corresponding proportion of total emissions they represent in the distribution, and their static follow-up thresholds. For example, to target the top 2% of leaks, static thresholds of 0.34 and 0.54 $g \cdot s^{-1}$ are required for empirical distributions A and B, respectively (inputs are described in the following section). Although we define static thresholds using a leak-size distribution, facility-scale fugitive emissions measurements could also be used when available. The MVL methods modeled in this study measure at the facility scale, which may aggregate multiple leaks. Therefore, a static threshold of 2% does not suggest that 2% of facilities will receive follow-up – far more than 2% of facilities will be flagged as there are often multiple leaks and design emissions. We use facility-scale emissions as a starting proxy for facilities with anomalously high leak rates. Facilities are flagged if multiple leaks are present and their summed emission rate exceeds the follow-up threshold.

Dynamic thresholds depend on the relative emissions among facilities at the time of screening. During a survey, a screening method visits each facility in a program, quantifies the emissions of each facility, ranks facilities according to emission rate, and sends a follow-up crew to a specified proportion of the highest-emitting facilities (e.g., top 10%). The dynamic target defines what proportion of the highest emitters should be flagged for follow-up (the dynamic target is in this case is 0.1). Note that the dynamic threshold that follows from the dynamic target is implicit and does not need to be calculated. Compared to static thresholds, dynamic thresholds are conceptually simpler but may not be realistic for all screening methods. An advantage of dynamic follow-up is that program cost should be relatively constant, as the same number of facilities always receive follow-up. However, if aggregate program emissions are very high, dynamic thresholds are not guaranteed to meet mitigation goals relative to a baseline. Similarly, if emissions are lower than expected, dynamic thresholds may lead to unnecessary follow-up.

Table 5.1 Static targets and corresponding static thresholds used in this study. Two leak-size distributions are used; D1: Clearstone Engineering Ltd. (2018), and D2: Ravikumar et al. (2020).

Static target	Emissions proportion		Static follow-up threshold ($g \cdot s^{-1}$)	
	D_1	D_2	D_1	D_2
0.01	0.47	0.17	0.85	0.35
0.02	0.54	0.30	0.54	0.34
0.04	0.63	0.44	0.31	0.14
0.08	0.72	0.57	0.14	0.06
0.16	0.81	0.70	0.09	0.03
0.24	0.87	0.78	0.06	0.02
0.32	0.91	0.83	0.04	0.02
0.40	0.94	0.87	0.03	0.01
0.48	0.96	0.91	0.02	0.01
0.64	0.98	0.96	0.01	0.01
0.80	0.99	0.98	0.01	0.00
0.92	1.00	0.99	0.00	0.00

5.3.2 Equivalence scenario modeling

The parameters and empirical inputs used in this study are based on a LDAR-Sim case-study demonstration for Alberta, Canada (Fox et al., in review). Alberta produces approximately 0.3 billion m^3 of marketable natural gas and 0.5 million m^3 of crude oil per day from bituminous sands and a network of ~176 thousand conventional and unconventional oil and gas wells (AER, 2020, 2019). As of January 2020, LDAR must be performed at tens of thousands of facilities up to three times per year using OGI cameras or OVAs (AER, 2018; Johnson and Tyner, 2020). These new regulations are also among the first globally that allow producers to develop and implement 'alternative' LDAR programs, which can consist of any combination of technologies and work practices that demonstrate equivalence.

We establish equivalence scenarios using two empirical leak-size distributions: (D_1) the Clearstone Engineering dataset used in Fox et al. (in review), and (D_2) new data from Alberta (Ravikumar et al., 2020). In D_2, we use only leaks from initial LDAR surveys, as these are most representative of pre-LDAR conditions. Compared to D_1, D_2 is larger (n = 969 vs. 281) and better matches distributions seen elsewhere, where typically 5% of leaks represent 50% of emissions

(Table 4.1; Brandt et al., 2016). Only one company was surveyed in D_2, whereas D_1 represents 63 companies with broad geographical range.

Fugitive emissions are simulated under a range of follow-up targets (listed in Table 4.1). The same follow-up targets are used to establish static and dynamic thresholds. All simulations are run on 500 randomly selected facilities in Alberta. Locations of facilities do not matter to this simulation because weather and travel distances are ignored to generalize the study, but these parameters are important in applied situations. Similarly, method-specific parameters (e.g., detection limits, reporting and repair delays) and labour availability are not required to identify equivalence scenarios. For each static or dynamic target, 10 simulations are run over 6 years (our analysis excludes year 1). Following previous studies, we use a leak production rate (LPR) of 0.0065, but sensitivity to LPR is explored later (Fox et al., in review; Kemp et al., 2016). A weak operator (also referred to as a 'null repair rate') is used to simulate ongoing maintenance.

Quantification is only required for technologies that use triaging – not close-range technologies used to detect and diagnose individual leaks. In some treatments, design emissions and screening quantification errors are introduced. Design emissions are introduced following the methodology outlined in Fox et al. (in review). Facility-scale quantification error (E) remains poorly constrained for LDAR screening methods, and likely depends on the work practice used, dispersion modeling approaches, and a range of method-specific environmental factors. For example, quantification uncertainties for ground vehicles are reported to range from 50-350% (Fox et al. 2019). More recently, blind controlled release experiments found quantification estimates from mobile LDAR technologies to be within a factor of two ~35% of the time, and within an order or magnitude ~82% of the time (Ravikumar et al., 2019). Rather than attempt to replicate E for any given method, we present three hypothetical scenarios. In E_1, facility-scale screening quantification has zero uncertainty. In E_2, an error term is drawn from a normal distribution with a mean of 0 and a standard deviation of 2.2, such that ~35% of observations fall within ±1 of the true value (factor of two). In E_3, the error term is drawn from a normal distribution with a mean of 0 and a standard deviation of 7.5, such that ~82% of observations fall within ±10 times the true value (order of magnitude). In both cases, the error term is a percent change from the true emission rate to the estimated rate. In E_1, the true rate equals the estimated rate, and all follow-up decisions are optimal. In E_2 and E_3, discrepancies between the two rates reduce follow-up effectiveness as facilities are mis-ranked.

$$Q_E = \begin{cases} Q_T + Q_T \cdot E, & if\ E \geq 0 \\ Q_T/|E-1|, & if\ E < 0 \end{cases}$$

$$E = \begin{cases} 1, & if\ E = E_1 \\ Normal(0, 2.2), & if\ E = E_2 \\ Normal(0, 7.5), & if\ E = E_3 \end{cases}$$

5.3.3 Cost-effectiveness modeling

A balance exists between regulators, who require equivalent mitigation outcomes or better, and producers, who generally want to achieve equivalence at the lowest cost. If we assume that producers face no additional incentives to reduce emissions below compliance levels and that additional emissions reductions will cost more, producers will implement equivalence scenarios in practice. For each survey frequency, we extract optimum follow-up targets by linearly interpolating between the nearest programs above and below simulated regulatory emissions (illustrated in Figure 5.1).

Equivalent programs must offer cost savings over the regulatory program to be adopted by the O&G industry. However, estimating the cost-effectiveness of screening programs is challenging as the market is nascent and well-established workflows have not been rigorously demonstrated. Many companies exist, but each has a unique product or service that combines a platform (e.g., aircraft, drone, satellite, vehicle), its sensors, and some work practice. Costs are likely to change as these companies continue to develop and identify their role in the market. One way to estimate costs is to simulate typical screening programs in LDAR-Sim by platform class. However, parameterization of these programs is difficult, as little is known about daily costs, the number of facilities possible to screen per day, and general performance metrics of any specific solution.

Instead of forward simulating the cost of screening programs, we examine the amount of money that would be available for a screening technology within an over-arching program budget. This addresses the question of 'How much do screening programs need to cost to result in savings relative to a regulatory program?' This inverse question is easier to answer as fewer assumptions about screening effectiveness are required.

For any equivalence scenario, we know: (1) the regulatory OGI survey frequency (F_{reg}), (2) the dynamic target proportion (τ), and (3) the screening survey frequency (F_{screen}). If we assume that follow-up OGI inspections are identical to regulatory OGI inspections, the per-facility cost of OGI (C_{OGI}) should also be the same. Equivalent cost occurs when the cost of regulatory OGI (left side of equation) equals those costs of both screening and follow-up (right side of equation):

$$F_{reg} \cdot C_{OGI} \cdot (1-\varphi) = F_{screen} \cdot C_{screen} + C_{OGI} \cdot \tau$$

where φ is the discount in program costs expected. Solving for C_{screen}, the money available for screening after accounting for cost of follow-up is:

$$C_{screen} = \frac{F_{reg} \cdot C_{OGI} \cdot (1-\varphi) - C_{OGI} \cdot \tau}{F_{screen}}$$

To simulate the risk and costs of new program development, we use a discount requirement (φ). The discount can be used to calculate C_{screen} for screening programs that must be less expensive than the regulatory program. For example, if a producer requires the MVL program to be 30% cheaper than regulatory OGI to justify the effort and risk of adoption, φ = 0.3 is used to calculate C_{screen}. For cost equivalence, φ = 0. Average OGI costs are generally between USD $250 and $450 per facility (Environmental Defense Fund, 2016). The C_{screen} calculated here is conceptualized as the money available for per facility screening, or a contract bid price that a producer could offer to screening solution providers. If a screening technology cannot perform screening at this price, it is not cost effective. If a screening technology can perform the screening for less than C_{screen} additional cost savings are available in the program.

To illustrate C_{screen}, consider an MVL program that applies screening two times per year with τ = 0.1 (i.e., the top 10% of emitting facilities are flagged and receive follow-up after each round of screening). To keep the scenario simple, we exclude confounding aspects of design emissions and quantification error. The regulatory base case consists of 500 facility surveys using OGI with a total cost between USD $125,000 and $225,000. The MVL program would result in 1,000 facility screening surveys, with follow-up OGI inspection at 100 facilities. The OGI cost for 100 facilities is between USD $25,000 and $45,000. The money available for screening, C_{screen}, is between USD $100 and $180/facility when φ = 0, or between USD $62.5 and $112.5/facility when φ = 0.3. It is important to note that C_{screen} < 0 in situations where τ and φ approach 1. In other words, screening solution providers must work for free or even pay operators to perform the service – a situation that is unlikely to ever occur.

5.4 Results

5.4.1 Equivalence scenarios

Each screening and regulatory LDAR program results in a fugitive emission rate averaged over the 5-year simulation. Figure 5.1 provides an example of how increasing either the number of screening surveys or the amount of follow-up (τ) can lead to lower fugitive emissions. Increasing either screening surveys or τ results in diminishing returns in emissions reductions. Emissions reduction equivalence occurs when simulated emissions for a regulatory program (e.g., annual OGI) equal the emissions of a screening program. In theory, when the screening and regulatory programs have the same survey frequency, τ must equal 1 to achieve equivalence (i.e., all facilities will receive follow-up and screening becomes redundant). Therefore, each screening survey frequency above the prescribed regulatory frequency will have a corresponding τ required for equivalence. For example, if the regulatory program is triannual OGI at all facilities, screening once or twice per year will not be equivalent, and screening three times per year will only be equivalent when 100% of screened facilities receive follow-up ($\tau = 1$). The only sensible equivalence scenarios occur when $\tau < 1$, which requires a higher screening survey frequency than the regulatory program. As screening survey frequency increases, the corresponding τ required for equivalence decreases.

The simplest definition of an equivalence scenario is a screening survey frequency and corresponding dynamic (τ) or static target under a set of modeling assumptions. All equivalence scenarios developed in this study are shown in Figure 2. In panels A and B, equivalence scenarios closer to the top-right require more work (i.e., higher screening frequency and follow-up) relative to those in the bottom left. Achieving equivalence with a triannual OGI program requires either more screening or more follow-up than achieving equivalence with annual OGI. A similar pattern exists between programs with and without design emissions. Here, screening technologies are unable to differentiate fugitive from design emissions, meaning that facility-scale emissions estimates include both. The fugitive emission signal can therefore be lost in the noise of design emissions, which impacts facility ranking during triaging and results in follow-up crews being sent to the wrong facilities – ones with high design emissions but not necessarily high fugitive emissions. To achieve equivalence, the τ must increase to account for triaging errors, in some cases more than doubling the required number of follow-up inspections.

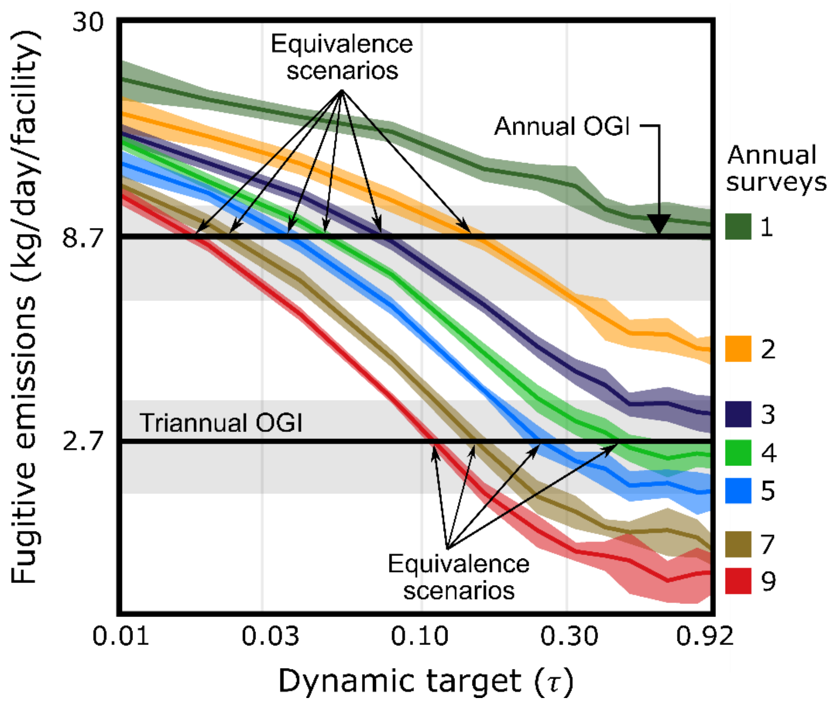

Figure 5.1 Ensemble of simulations across a range of dynamic follow-up targets and annual surveys. As the dynamic target increases, more facilities receive follow-up, more leaks are found, and fugitive emissions are lower. Increasing survey frequency similarly results in lower fugitive emissions as infrequent large emissions are found and repaired sooner. Intersections between regulatory OGI programs (black horizontal lines) and screening programs (colour lines) denote equivalence. For example, to be equivalent with triannual OGI, a screening program that screens all facilities 9 times per year must send OGI follow-up to the top 12% of highest-emitting facilities. Shaded regions are two standard deviations ensemble variability in emissions rate. This example shows simulations under D1, E1, and without design emissions (the most favourable situation for screening technologies to be economically viable). Fugitive emission rates are averaged over 5-year simulations.

Triaging is also impacted by quantification error. With no screening quantification error (E_1), facilities can be correctly ranked according to emission rate before follow-up inspectors are dispatched. As quantification error increases to E_2 and E_3, facilities are ranked not according to their true emissions rate (Q_T), but according to the estimated rate (Q_E), effectively 'shuffling' the ranking. Whereas the impact of design emissions is relatively consistent for different survey frequencies, quantification error is more problematic for MVL programs with lower screening survey frequencies. This occurs because quantification errors are a percentage of the true emission rate, whereas design emissions are drawn from an empirical distribution and are assumed independent of the fugitive emissions at a facility. When screening survey frequency is high, less time passes between surveys, which allows less time for leaks to accumulate. When Q_T is close to zero, high quantification errors have less absolute impact on Q_E. Future work should investigate Q_E should be estimated as a percentage of Q_T or as an absolute difference.

Comparing panels, A and B in Figure 5.2 illustrates the impact of leak-size distribution. Equivalence scenarios are achieved with less work for screening programs under D_1, which is more heavily skewed (Figure 2A). When a smaller number of large leaks comprise the majority of total emissions, strategies that focus on these leaks benefit. Figure 2B shows equivalence scenarios under the D_2 leak-rate distribution, which is less heavy-tailed and consequently requires a larger number of follow-up inspections to achieve equivalence.

Figure 2C shows equivalence scenarios using static targets, and must be interpreted differently from panels A and B. Both design emissions and quantification error result in lower required targets. When design emissions are added to fugitive emissions at a facility, it makes it easier for the static threshold to be surpassed. However, this results in much more work, which is not immediately evident as a higher static target does not correspond intuitively with follow-up requirements like it does with a dynamic target. Lower static targets are required to achieve equivalence as quantification error increases due to the heavy-tailed shape of the facility-scale emissions distribution. More facilities will fall below a given follow-up threshold than above, meaning that quantification error will push lower-emitting facilities above the threshold more often than higher-emitting facilities below the threshold. This results in more work being done than necessary, leading to a lower static target.

Previous studies have shown the influence of LPR on simulation results (Kemp et al., 2016; Fox et al., in review). We present simulations under a range of LPRs to determine whether results in this study are sensitive to LPR (Figure S1). Fugitive emissions vary greatly between programs, but because emissions increase for both the regulatory program and the screening program, equivalence scenarios are robust to differences in LPR.

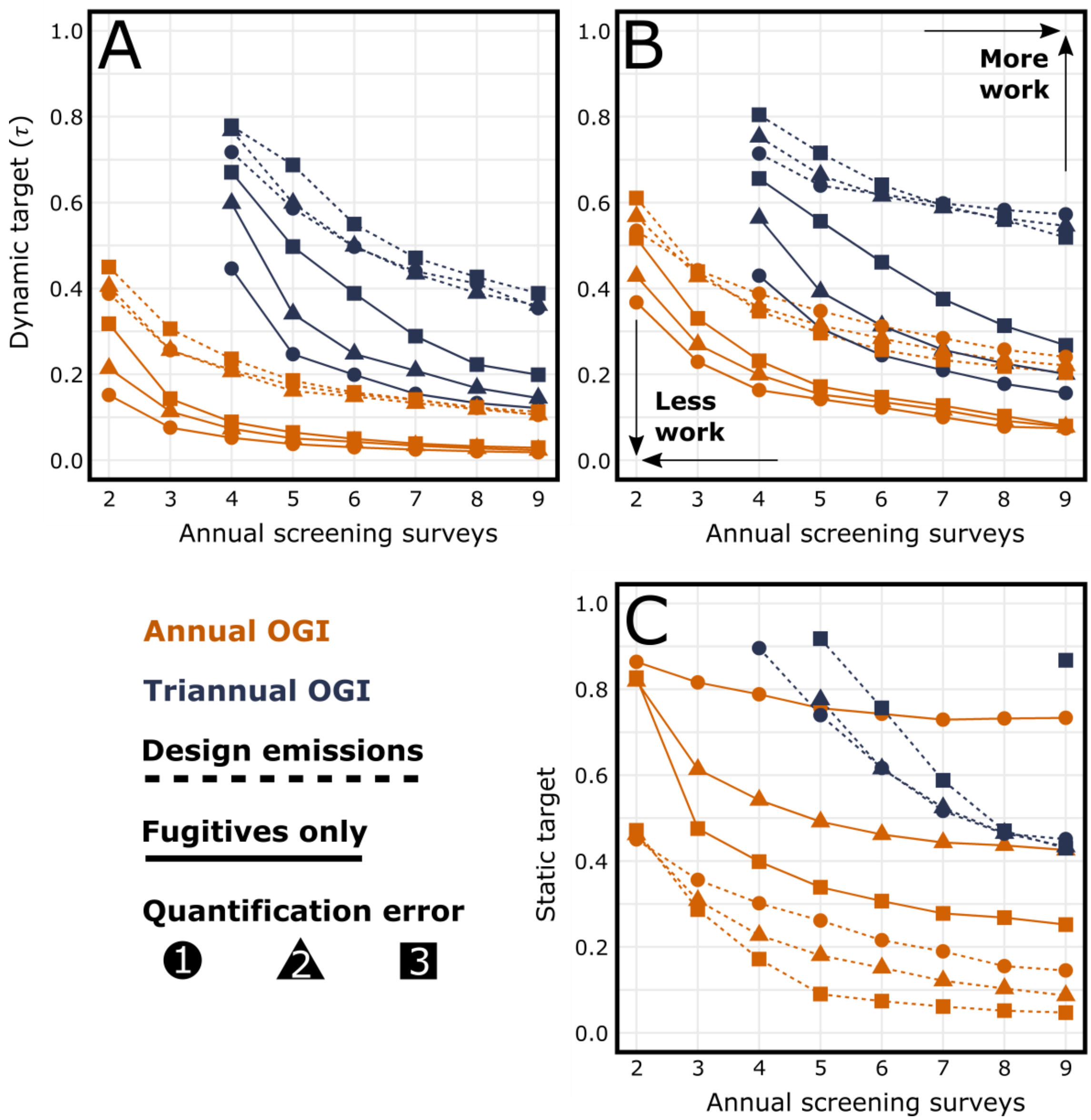

Figure 5.2 Catalogue of equivalence scenarios under various conditions: (A) Dynamic targets assuming leak-rate distribution D1 (more skewed), (B) Dynamic targets assuming leak-rate distribution D2 (less skewed), and (C) Static targets assuming D1. Each point represents an equivalence scenario with either annual (orange) or triannual (blue) conventional OGI programs. Solid lines optimistically assume that all emissions are fugitives, while dashes lines assume the presence of design emissions. Quantification errors increase from E1 (perfect facility-scale quantification) to E3 (order or magnitude error) and are explained in the main text. Note the different interpretations of dynamic targets (τ; proportion of top-emitting facilities requiring follow-up) and static targets (proportion of the fugitive leak-size distribution to target).

5.4.2 Cost-effectiveness

In Figure 3, we estimate C_{screen} for φ = 0 (blue) and 0.3 (orange) where C_{OGI} = USD $250/facility (lower limit) and $450/facility (upper limit). Under best-case scenarios for MVL, which assume no design emissions, E_1, and D_1 (extreme skew), C_{screen} rarely exceeds $100/facility. To reduce costs relative to the regulatory program by 30%, screening costs must not exceed USD ~$50–$90/facility. The worst-case scenario is more realistic; it accounts for design emissions, quantification error (E_3), and assumes D_2, which better approximates emissions elsewhere (Brandt et al., 2016). Every screening program under the worst-case scenario is more expensive (i.e., negative values) than regulatory OGI before screening is paid for. In other words, triaging is so inefficient that, to achieve equivalence, more follow-up surveys must be conducted than would be required under the exhaustive but less frequent regulatory program. To illustrate how C_{screen} can be negative, consider a hypothetical LDAR program with 100 facilities. A regulatory program that requires annual OGI will require 100 inspections. A proposed screening program with biannual surveys will have $C_{screen} = 0$ if 50 follow-up surveys are required after each survey ($\tau = 0.5$), as the total number of follow-up surveys (100) equals the number of required OGI surveys under the regulatory program. Therefore, C_{screen} for the biannual screening program can only be positive when $\tau < 0.5$. If screening quantification error reduces triaging effectiveness by introducing triaging errors, τ may increase above 0.5, causing C_{screen} to drop below zero.

In general, the screening programs simulated in this study are more cost-effective when: (1) C_{OGI} is higher; (2) equivalence scenarios with fewer screening surveys are used (i.e., those with higher τ and more follow-up); (3) when design emissions are lower; (4) when quantification error is low; and (5) when heavier leak-size distributions are assumed. Results also suggest that more money is available for screening when the regulatory program requires more OGI surveys. This is likely due to the relative burden of going from one to two surveys being greater than the burden of going from three to four.

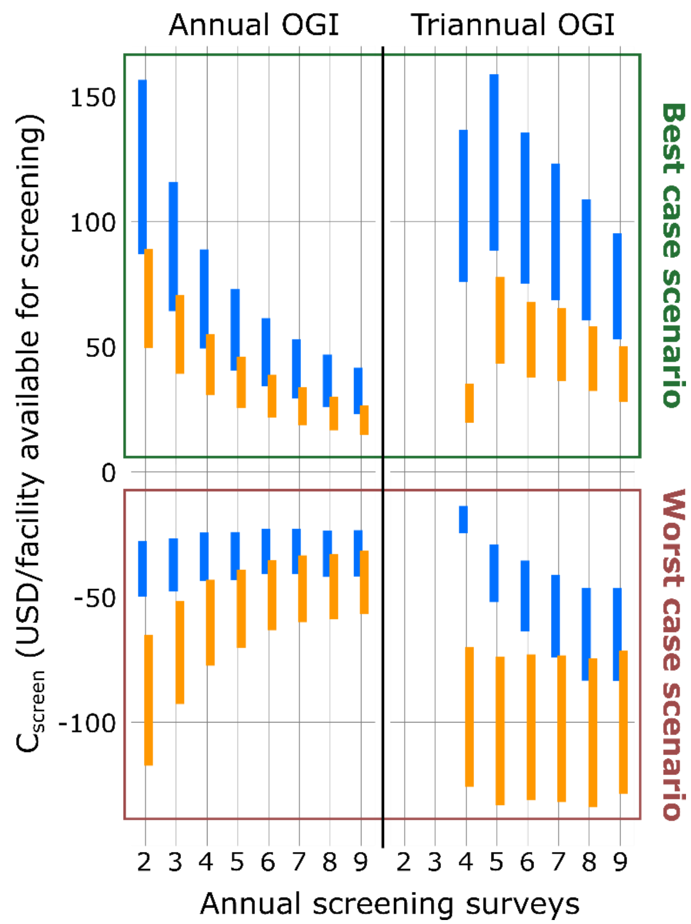

Figure 5.3 Maximum screening program costs (C_screen) under a range of equivalence scenarios. Estimates are made assuming OGI survey costs of USD $250/facility (lower limit of each line) and USD $450/facility (upper limit of each line). Blue lines show required screening costs when screening program target costs are equal to those of the regulatory program (φ = 0). Orange lines show required screening costs for screening programs that offer a 30% cost reduction relative to the OGI-only regulatory standard (φ = 0.3). The best-case scenario for MVL assumes perfect screening quantification (E1), leak rates from D1, and no design emissions. The worst-case scenario assumes E3, D2, and presence of design emissions. Negative C_screen values indicate that screening technology providers must pay to perform services., a situation unlikely to exist in practice.

5.5 Discussion

Our results suggest that MVL methods may struggle to be cost-effective compared to regulatory OGI (Figure 3). Design emissions and quantification errors can lead to incorrect ranking of facilities by emission rate, which increases the required number of follow-up inspections. Simulations suggest that MVL will be more cost-effective if triaging errors can be reduced. While negative C_{screen} values are clearly unprofitable (screening providers would have to pay to provide their service), it is difficult to infer profitability when screening must be $50 - $100. Screening technologies are variable in approach as the industry is nascent and is repurposing technologies used outside the O&G industry. For example, satellites are expensive to build, launch, and operate, but are not constrained by access to services such as lodging for crews and road access. Vehicle-based systems are constrained by driving time but are more scalable as the hardware and deployment is less expensive than satellites. These two examples are ends of a spectrum; many variables affect the cost of screening surveys, which may be unprofitable in some contexts but could be successful elsewhere.

The scenario with parameters least favourable for MVL viability ('worst-case scenario') considered in this study may be optimistic due to five modeling assumptions. First, our analysis does not consider the presence of false positive and false negative detections for screening technologies, which could introduce additional triaging errors. False positive and false negative detections exist among screening methods but are more common when the emission rate is close to zero (Ravikumar et al., 2019). False negatives are less likely to occur with the largest and most consequential emission sources as many detection probability increases with emissions rate (Ravikumar et al., 2019). Second, simulations account for temporal variability of design emissions, but not of fugitive emissions. If fugitive emissions are episodic, then problems arise for MVL including (1) detecting an episodic fugitive and dispatching a follow-up inspector who is unable to identify the source, and (2) failing to detect an episodic source that was not emitting at the time of screening. For MVL to be effective, the facilities that receive follow-up must be the highest-emitting facilities over time and must be clearly distinguishable from low-emission facilities. Given that screening is a snapshot in time, the implicit assumption is that instantaneous emissions are representative of the long-term average, which may not be the case (Johnson et al., 2019). Third, follow-up OGI could be more expensive than regulatory OGI, because flagged facilities may be spaced further apart, which increases travel time and cost. Fourth, this study assumes that screening surveys are spaced evenly apart throughout the year and that they aren't impacted by adverse environmental conditions or logistical constraints. However, all screening methods have specific operational limitations that may further impact performance, such as clouds for satellites, roads for vehicle systems, snow for LiDAR, and wind and precipitation for drones (Fox et al., 2019a). Fifth, we assume that screening technologies have minimum detection limits sufficient to accurately rank at least as many facilities as require follow-up. However, the most cost-effective scenarios require a small number of screening surveys with considerable follow-up (τ up to 0.8 for triannual OGI equivalence. When τ is higher, the corresponding minimum detection limit required by the screening method is lower.

Conditions favourable to screening may exist. Our results are sensitive to the skewness of leak-size distributions, with heavier tails favouring MVL. Lower design emissions also favour MVL. Facility-scale design emissions will likely decline over the coming decade as regulators and operators work towards mitigation targets. Clear steps can be taken to reduce design emissions, such as conversion from high- to low-bleed pneumatics, installation of vapour recovery units on storage vessels, and best practices for well completions and manual liquid unloadings.

Reducing quantification error for screening technologies may also improve MVL but is a major challenge. Atmospheric variability is a significant source of error (Caulton et al., 2018). Some mobile methods have reduced quantification error by increasing measurement times (Brantley et al., 2014). However, increased measurement times typically come with a cost penalty. Additional work will be required to understand whether the trade-off between measurement time and quantification error can be sufficiently reconciled to enable cost-effective screening. Possible workarounds may exist. For example, low-sensitivity screening methods may opt for $\tau = 1$, eliminating the need for triaging and thus avoiding incorrect ranking. However, minimum detection limits must be consistent for screening methods to ensure that all facilities emitting above a specified amount receive follow-up.

It may be possible to reduce triaging errors with better ranking algorithms. The simple static and dynamic thresholds used here are intuitive yet crude; more complex models may exist that incorporate facility-specific quantification errors and knowledge. For example, if estimated emissions exceed production, it may be possible to identify an error and require additional measurements. However, complex follow-up schemes that are not codified can be difficult to regulate. It is important to note that these results are interpreted in the context of LDAR. MVL technologies may also provide additional value beyond the immediate goal of mitigating fugitive emissions. For example, facility-scale emissions data, even if inaccurate, may be desirable for producers for reporting, tracking progress, or estimate probability of non-compliance with facility-scale venting limits. Some screening technologies can also produce ancillary data, such as high-resolution aerial imagery from aircraft. Screening can also be performed passively if measurement systems are deployed on existing vehicles.

Equivalent and cost-effective MVL programs may be elusive, but we remain in the earliest stages of understanding how these systems might contribute to methane monitoring and mitigation. Better performance will inevitably result from stricter venting limits, improved sensors and quantification algorithms, and a better understanding of how, where, and when to use MVL. It should be noted that this study only applies to facility-scale MVL, and that additional analyses are needed to evaluate equipment-scale screening, continuous monitoring, single-visit models, and other approaches to alternative LDAR.

5.6 Conclusion

This study used LDAR-Sim – an agent based numerical LDAR model – to explore scenarios involving screening technologies. Our results suggest that (1) emissions-reduction equivalence is possible between regulatory OGI programs and MVL, and (2) cost-effective equivalence scenarios exist but may be difficult to achieve. Circumstances that impact performance also impact cost-effectiveness, as more surveys are required to achieve equivalence. We have shown that these circumstances include design emissions, quantification error, and the empirical leak-size distribution used for modeling.

Our simulations suggest the largest impediment to cost-effective MVL is the confounding presence of design emissions. Placing tight regulatory focus on leaks only, instead of all methane emissions from a facility, may limit uptake of screening. The primary issue is scale – to find leaks on a facility with extensive design emissions requires fine-scale measurements, which favours close-range approaches. In terms of chemistry and environmental impact, emissions from design sources such as vents are usually identical to those from leaks. Facility-scale screening technologies could be more effective if the fugitive-design emissions dichotomy were replaced with a more general focus on identifying and resolving the highest-emitting facilities.

Chapter 6

Conclusion

6.1 Summary of contributions

My thesis was motivated by the need to evaluate current and emerging methane-sensing technologies to better understand the roles they can and should play in reducing emissions from the oil and gas industry. Specifically, I targeted three knowledge gaps:

1. What technologies exist and what is our current understanding of how they work?
2. How should emerging technologies be integrated into LDAR programs?
3. Can multi-visit LDAR (MVL) reduce emissions cost-effectively?

Knowledge gap 1 was targeted by the literature review presented in Chapter 2 (Fox et al., 2019a). A broad range of new methane-sensing technologies have been developed over recent decades, yet only a small number are used in LDAR programs. Most of the literature to date has focused on the use of new sensors and platforms for research that seeks to understand methane emissions from O&G (e.g., Karion et al., 2015; Nathan et al., 2015; Yacovitch et al., 2015). These technologies and studies have been used to inform, but not necessarily to mitigate. Our review is the first published work to explore and evaluate how handheld instruments, mobile ground labs (MGLs), fixed sensors, drones, aircraft, and satellites might be used not only for research, but as part of industry led LDAR programs.

We find that each technology class shows potential and that, because they are so diverse in strengths and limitations, each may contribute in different ways. Currently, handheld instruments, aircraft, and MGLs show the greatest potential. Handheld instruments may be labour-intensive, but they are established in government, have been shown to work well, and are familiar to industry. Aircraft and MGLs have been used extensively and with different types of sensors and analytics. They offer more rapid surveys, despite lower sensitivity, which may allow them to play a role in identifying the small number of large sources that account for a disproportionate share of total emissions. At the time of publication, aircraft and MGLs have not been used for LDAR. However, several alternative LDAR programs in Alberta are gearing up to use these platforms. Drones and satellites each face challenges at this time that may limit their uptake in the near term, but they may eventually play a role. Drones do not yet offer obvious labour reductions relative to handheld instruments, and are particularly vulnerable to adverse weather such as precipitation and strong wind. Fixed sensors show great potential, but costs and analytics must both improve. Satellites have been subject to considerable excitement, but have yet to transparently demonstrate reliable point-source detection and quantification performance.

Unlike handheld instruments, most of the new technologies available are unable to pinpoint individual leaks for diagnosis and repair. Drones, MGLs, fixed sensors, aircraft, and satellite all operate either by positioning themselves inside the methane plume downwind of the source, or by using remote sensing. In both cases, the information products generated do not have sufficient spatial resolution to understand the root cause at the component scale. Our review is the first published work to explicitly identify this limitation, and the first to introduce the concept of 'screening' as a solution. In screening, mobile technologies rapidly survey facilities to identify those with high emissions, and send inspectors with handheld instruments to follow-up and ultimately repair leaks. We introduce the concept of a 'comprehensive monitoring program' (CMP), which may integrate numerous different technologies at different spatial and temporal scales into a single LDAR program.

Some of the technologies promising to improve LDAR may be ready for oilfield, but others may fall short. Until recently, most LDAR regulations have mandated periodic inspections using handheld instruments to exhaustively survey all potentially leaking components. However, over the past few years, several forward-looking governments have introduced regulations that provide a means of entry for new technologies. In Alberta and elsewhere, these regulations say that any alternative LDAR program may be considered, but that it must be demonstrated to result in equivalent emissions as the standard regulatory program. However, demonstrating emissions equivalence among programs with vastly different technologies is fraught with complications, and governments have provided limited guidance on how to do so.

Knowledge gap 2 was the absence of a clear standard for effectively demonstrating equivalence. This challenge was addressed in Chapter 3, which set out a regulatory framework, and Chapter 4, which introduced a new simulation framework, both for demonstrating equivalence. The regulatory framework presented in Chapter 3 brought together diverse stakeholders from across North America in three workshops held in Calgary, Alberta and Fort Collins, Colorado. Over 100 participants contributed to the framework, representing federal governments, provincial and state governments, large, medium, and small producers, service providers, technology developers, academics and researchers, industry groups, environmental non-profits, and consultants. The resulting framework proposed using a sequential combination of controlled testing, simulation modeling, and field pilots to demonstrate equivalence. Briefly, controlled testing would be used to develop performance metrics, such as detection probability curves and false-positive rates under a range of environmental and operational conditions. Performance metrics could then be used in a simulation framework to estimate mitigation of an LDAR program that uses the proposed technology. Finally, if simulations suggest that the proposed alternative LDAR program will achieve equivalence, a pilot field deployment should occur on a subset of facilities to identify any unforeseen problems and to further evaluate mitigation effectiveness. Although the framework was widely approved by all stakeholder groups, several important challenges were identified. These included challenges with (1) ensuring representative controlled testing, (2) developing representative models that are both accessible and flexible, (3) understanding how new work practices, such as screening, should be regulated, especially when follow-up decision-making is required, (4) awareness of human factors such as incentivization and compliance, and (5) logistical considerations such as transparency, inter-jurisdictional cooperation, and oversight.

The open-source simulation framework presented in Chapter 4, called LDAR-Sim, was designed to enable transparent, flexible, and reproducible comparison of LDAR programs. LDAR-Sim can serve many different roles in addition to emissions reduction equivalence; (1) it can be used by producers to evaluate the cost-effectiveness of a proposed program, (2) it can be used by regulators to evaluate the emissions reductions that can be expected from different regulations, (3) it can be used by technology developers to better understand trade-offs among sensitivity, cost, and other performance metrics under different deployment scenarios, and (4) it can be used by researchers to test hypotheses surrounding LDAR, weather, technology, labour, costs, and work practices. The work in Chapter 4 reveals that demonstrating equivalence requires a clear and detailed understanding of regulatory and alternative programs, including performance metrics, environmental limitations, labour availability, leak production rates, and so on. There exist different ways to compare programs, and each may arrive at different conclusions. Chapter 4 also shows for the first time that vented emissions can reduce the effectiveness of screening technologies that are unable to differentiate between fugitive emissions and other sources.

Knowledge gap 3 was addressed in Chapter 5 by using LDAR-Sim to evaluate the promise of MVS. Work in Chapter 4 revealed the need to better understand the impact of confounding emissions sources on screening technologies that are required to target only fugitive emissions. For alternative LDAR programs (including MVS) to be adopted, they must be not only equivalent, but they must also be more cost-effective than the standard regulator program for industry to be interested. Simulations in Chapter 5 show that MVS can be equivalent or cost-effective, but rarely both at once. The most cost-effective programs exist for screening technologies with low quantification error and facilities with limited confounding emissions from venting or combustion. However, most current screening methods have very high quantification error and most current O&G facilities have design emissions that exceed fugitive emissions.

6.2 Implications and future work

My thesis shows that innovation is exploding in the O&G methane-sensing space, but that significant uncertainty exists in how to demonstrate, validate, regulate, and deploy emerging technologies, methods, and ideas. Rapid advances in technologies and methods for measuring methane emissions are outpacing capacity to fully understand how each one works. This momentum brings danger and opportunity. On one hand, when uncertainty is high, the risk of permitting and deploying ineffective solutions increases, which could limit emissions mitigation. On the other hand, a willingness to make mistakes should accelerate innovation and reduce uncertainty. Some jurisdictions, including Alberta, are following the latter path, and are currently approving alternative LDAR programs for piloting that may ultimately prove ineffective or too expensive. Other jurisdictions are watching and ready to learn from the mistakes that will inevitably be made. As third-party off-site monitoring capacity and prevalence develops (e.g., MethaneSAT), and as reporting practices and compliance efforts are honed, ineffective solutions will become evident. Similarly, the most cost-effective solutions with proven performance will grow in popularity. Although Chapter 5 suggests that much of the excitement surrounding screening technologies may be premature, both the innovation landscape and supporting regulations are nascent; great potential exists.

Significant additional research will be required to better understand how new technologies perform and how they might contribute to mitigating O&G methane emissions. A crucial step is to work towards thoroughly understanding individual technologies and methods. First, researchers should answer the question of how best to conduct evaluations and interpret the results. Standardized testing protocols that evaluate a broad range of relevant performance metrics should be developed for different technology classes. Controlled release testing under a range of environmental conditions is necessary for representative understanding. Work practices must be explicitly defined before testing. For example, aircraft that fly at 300 m during controlled testing must not fly at 600 m during piloting to increase survey speed, as the increased distance would lead to lower sensitivity. Similarly, systems tested under ideal weather should not be assumed effective in a snowstorm. Ideally, testing should be performed by a reputable third party, and protocol development should be led by independent researchers in consultation with regulators, producers, innovators, and service providers. To be adopted, results, methods, and testing conditions should be made publicly available. Technologies and methods must also be piloted to ensure that they are effective in the real world. Many technology developers and innovators are unfamiliar with the oilfield, and are not aware of specific logistical challenges such as obtaining site access, safety protocols, and how to work within existing operational structures and constraints.

Understanding LDAR programs is just as important as understanding the technologies and methods they use. More research is needed to reveal the impacts of different survey frequencies, minimum visitation intervals, repair practices, reporting, program evaluation, and so on. Perhaps most importantly, new technologies, which are often evaluated on the basis of whether they can detect emissions, should be evaluated based on the broader context of how they might contribute to mitigation. Many emerging solutions have focused a disproportionate amount of energy to determine how to achieve detection and quantification, but have largely ignored the broader challenge of what to do with this information. Developing clear plans for what to do with measured data is crucial because measurements are not useful unless they lead to repair or resolution of emissions sources. Specifically, follow-up decision-making for screening methods and repair practices must be explicitly defined and take into account design emissions, quantification error, leak lifecycles, environment, and other method-specific limitations. More research is needed to understand the relative impact of these and other variables that may not yet be considered in LDAR-Sim and other models.

At first, new technologies and methods will be deployed individually and gradually as part of pilot programs. Eventually, efforts should be made to develop and test the comprehensive monitoring programs introduced in Chapter 2, integrating multiple measurement systems into a single LDAR program. Comprehensive LDAR programs will take into account the limitations and constraints of different measurement systems. Technologies and methods will be combined according to their strengths in the context of program facility types, environment, site densities, infrastructure availability, and so on. Designing and testing comprehensive programs can begin at research institutions and should leverage available simulation tools. Modeling as a means of understanding and comparing LDAR programs must also be improved. This means conducting more research to establish the 'baseline' data that underlie modeling assumptions. Most importantly, the 'leak lifecycle' of each region or company must be adequately constrained. Knowing leak production rates, null repair rates, emissions distributions, temporal variability of sources, and super-emitter characteristics for the assets of interest is crucial for developing programs that can be effective. For example, O&G producers with high null repair rates (due perhaps to strong best management practices) will limit the efficacy of screening technologies, especially if large leaks are generally found and resolved quickly by site operators. Without context, screening technologies may appear effective at finding large leaks, but they may be redundantly identifying the super-emitters that attentive operators are already able to identify on their own.

The human dimensions of LDAR were not studied in this thesis and have not been addressed elsewhere. Incentivization structures at all scales must be clearly understood by regulators and industry. Academic research can contribute to an improved understanding of what motivates various agents (e.g., inspectors, operators, CEOs, compliance officers) to contribute to methane reduction efforts. Different technologies, work practices, and even regulations may unknowingly offer opportunities for corner-cutting or avoiding compliance. Many such possibilities exist. For example, inspectors may be incentivized to conduct surveys as quickly as possible, and may intentionally or unintentionally overlook certain leaks. Without sufficient regulatory oversight, producers may be incentivized to adopt technologies that save money at the expense of emissions mitigation. Given the intensely competitive market for emerging technologies, strong incentive may exist to oversell performance capabilities. Given the steep technical learning curve for understanding LDAR effectiveness, producers may not know how to choose the right solutions. Education will therefore remain a crucial ingredient for success. Independent measurement-based research by academia and non-profits should continue as it may help to hold emitters accountable.

Perhaps most importantly, more fundamental research is imminently required to better constrain the methane challenge. Most basic questions remain either partially or entirely unanswered. For example, how do leak-size distributions, leak-count distributions, leak causes, super-emitter incidence, and temporal variability in emissions sources vary by facility type, equipment type, operator, facility age, production type, and so on? What explains the continued discrepancy between top-down measurements and inventory estimates? Considerable investments have been made into the deployment of new technologies without a clear understanding of the problem they are trying to address. A more effective path forward would be to increase research funding to better constraint the methane challenge, and then look for the best solutions to that challenge or develop new solutions if needed.

The current policy landscape for reducing O&G methane emissions is progressive and engaged but nascent. I am proud and encouraged that Alberta has adopted a version of the equivalence framework presented in Chapter 3. Looking to the future, Alberta and other jurisdictions can enable further innovation by assessing the value of separately regulating fugitive and vented emissions and providing specific reference standards and modeling assumptions for demonstrating program equivalence. As regulators around the world continue to improve and implement methane mitigation policy, they should strive to accelerate learnings through effective and transparent reporting and data sharing. However, much more work is needed to understand which policies are effective and why. Entire fields of policy scenario research remain mostly untapped, including developing and evaluating innovative compliance efforts and understanding the long-term environmental, economic, and innovation trajectories of different policies.

6.3 Final thoughts

When I began my doctorate in 2016, very little was known about emerging methane-sensing technologies and whether they might have a role to play in LDAR. At the time, words and concepts like ‘screening’ and ‘follow-up’ didn’t exist, and even if they did, regulations in Alberta and Canada did not yet enable the deployment of mobile technologies or the implementation of alternative LDAR programs. Today, the first alternative LDAR programs are being approved for use in Alberta, and regulators across Canada and the US are working towards progressive policies that foster innovation by enabling demonstration of emissions reduction equivalence. Innovation, demonstration, and a better understanding of LDAR have blossomed. Moving forward, we must continue to foster innovation and welcome new ideas while retaining a healthy skepticism of the ability of emerging solutions to achieve what they promise. I’m excited to continue working in this space and to look back on this thesis in 5, 10, and 20 years to see how far we’ve come.

Chapter 7

Appendix A: Supplementary information for Chapter 4

7.1 Operator detection

In the past, LDAR simulation modeling has operated on the assumption that the input emissions data represents a ‘pre-LDAR’ steady-state baseline, which must ensure that (1) system-total emission rate is in a steady state, (2) system-total number of leaks is in a steady state, and therefore (3) the relative proportion of leak sizes is in a steady state (Kemp et al., 2016; Ravikumar and Brandt, 2017). In the absence of LDAR, two variables define the leak pool. The first is LPR, which increases the number of leaks. The second, opposing process is the Null Repair Rate (NRR), which removes leaks to place an upper bound on the size of the leak pool. The NRR represents periodic inspection, maintenance, and repair by operators outside of an LDAR program. Together, LPR and NRR work in opposition to maintain the equilibrated baseline in the absence of LDAR.

To maintain a steady state, leak removal must equal production over time, both in terms of counts and total emissions. In LDAR-Sim, an ‘operator’ agent is used for null repair. The operator module is optional and can work both in the presence and absence of an LDAR program. At a specified interval (typically on Mondays), an operator visits each site (i.e. facility) in the simulation. At each site, each leak has a detection probability (NRR) at time t of:

$$NRR = \frac{LPR \cdot \mathrm{V}}{\overline{L_{t0}}} \cdot \frac{L_t}{L_{t0}} + \frac{p_{max} \cdot Q_i}{Q_{max}}$$

where V is the visitation interval in days (in this case, 7), L_{t0} is the global leak count at initialization, L_t is the number of active leaks at time t, and $\overline{L_{t0}}$ is the mean number of leaks per site at initialization. Note that units differ for NRR ($p \cdot leak^{-1} \cdot day^{-1}$), which is applied to individual leaks, and LPR ($p \cdot site^{-1} \cdot day^{-1}$), which is applied at the site level. We therefore scale LPR to the leak level using $\overline{L_{t0}}$. The unitless term $\frac{L_t}{L_{t0}}$ is used to maintain the steady state, adjusting NRR in proportion to the number of active leaks relative to initialization.

The final term is used in SA to explore the possibility that larger leaks are more easily identified and readily repaired by facility operators. In FEAST, LPR is applied to update the Markov state of non-leaking components. These components can be thought of as leak 'slots' that get filled with random draws from the leak-size distribution and then emptied via null repair or LDAR. If a big leak is drawn, the leak will probably be detected during the next LDAR survey and removed from the slot. However, if the leak is very small and unlikely to be detected, it will occupy that component. As time goes on, a greater proportion of the leak slots can become filled with undetectable leaks, unless they are randomly removed from the leak pool via the null repair rate. FEAST's Markov chain therefore requires that small leaks be equally likely as large leaks to be repaired in the null scenario. Otherwise, small leaks would grow to dominate the available components, 'taking up space'. However, it is reasonable to assume that larger leaks are preferentially removed from the leak pool during null repair because (1) large leaks are more easily detected by human audial, visual, and olfactory (AVO) senses, and (2) if discovered, large leaks are of greater safety and environmental concern.

Changing the NRR to be a function of leak size is challenging because it can destabilize the equilibrium of the leak pool. If we assume that removal rate is a function of leak size, the production rate must also change to keep the leak-size distribution in equilibrium. In other words, if larger leaks are more likely to be removed, then production of larger leaks must increase to maintain the relative proportion of large leaks, and production of small leaks must decrease to maintain the relative proportion of small leaks. Ideally, LPR and NRR should be derived independently. However, measuring these values independently is challenging (if not impossible), as NRR is nearly always implicitly measured at the time LPR is measured. One solution is to measure NRR and subtract it from LPR. However, measuring NRR may not be as simple as asking operators to keep track of each leak they actively repair; leaks may resolve themselves through changes in pressure and production, the emergence of different leaks nearby, or unintentionally during maintenance or changes in equipment. Furthermore, LPR and NRR likely vary by producer, basin, and facility.

In LDAR-Sim, we adopt a simple method of adjusting the operator detection function. This method ignores LPR-NRR dependence and is used only in the SA for heuristic purposes. A sigmoid-shape curve similar to those used for modeling detection limits for other methods could be used (Ravikumar et al., 2018). However, parameterizing the model is not currently possible, as there exists no empirical evidence describing AVO detection probability as a function of leak size. Here, p_{max} (called 'operator bonus' in the SA) is the marginal increase in detection probability assigned to the largest source in the empirical dataset, Q_i is the emission rate of the leak under consideration, and Q_{max} is the largest maximum emission rate in the dataset. In this way, each leak is assigned the same base probability of detection in addition to a 'bonus' chance of detection, which is a function of emission rate. Leaks emitting at the highest possible emission rate, as defined by the empirical leak-size distribution, will receive a bonus detection probability of p_{max}. For the smallest leaks, $\frac{Q_i}{Q_{max}}$ approaches zero, so the adjustment is negligible (Figure 7.3).

This approach makes two assumptions. First, detection probabilities increase with the size of the leak, up to the user-defined maximum. However, detection probabilities are not similarly or opposingly reduced for small leaks. Although it is reasonable to assume that large leaks will be quickly identified by operators (for an extreme example, recall the Aliso Canyon leak in northern Los Angeles, US during 2015-2016), small leaks may continue to be repaired at the base level due to periodic maintenance. Second, this approach assumes that LPR remains constant, meaning the baseline will necessarily disequilibriate as 1) large leaks are preferentially detected and repaired, and 2) more leaks are detected in total. Table 7.3 provides a simple demonstration that making NRR a function of leak size can have a large effect on mitigation, even for small p_{max} values.

7.2 Case study parameterizations

All empirical inputs used are specific to Alberta. Leak-rate and leak-count distributions come from 2017 data targeting 333 upstream O&G sites using OGI cameras (Clearstone Engineering Ltd., 2018). The Clearstone study was broadly representative of Alberta's O&G sector, involving 63 companies producing natural gas and light, medium, and heavy crude oil. Site-level measurements used to derive empirical vents were acquired using a combination of tracer flux and Gaussian dispersion methods (Zavala-Araiza et al., 2018).

We estimate LPR using publicly available LDAR data compiled from nine producers in Alberta between July 2007 and October 2016 (GreenPath Energy, 2017). In total, there were 8097 leaks measured, mostly at larger facilities such as gas plants and multi-well batteries. Of 1283 unique site visits, 436 were documented repeat visits of a minimum 30-day survey interval. We estimated LPR by dividing leak count by the number of days since the most recent survey. We removed one extreme outlier in which 78 leaks were detected after only 59 days. The distribution of site-level LPR estimates from the GreenPath study is shown in Figure 7.4 (note log scale). In FEAST, a constant LPR of 1e-5 per component is used. As FEAST assumes 650 components per site, the site-level LPR is ~0.0065 (vertical red line in Figure 7.4). Differences between FEAST and our estimate could be due to infrastructure types, region, methodological discrepancies in data acquisition, and validity of assumptions. Specifically, both studies assume that (1) leaks reported during the first survey were repaired, and (2) NRR is inconsequential. We cannot comment on the data used to estimate LPR in FEAST, but the data used in this study are highly unlikely to meet the assumption of complete repair. Repair rates for the dataset may be as low as 8%, due to incomplete record keeping and flexibility to defer repairs at the time of the study, which could lead to a significant overestimation of LPR.(GreenPath Energy, 2017) Nevertheless, LPR values estimated in this study span over two orders of magnitude. In FEAST, LPR was the most sensitive parameter, despite using only a limited range in the SA (the shaded region in Figure 7.4). For this demonstration we adopt the same LPR as FEAST (0.0065) but note that both the methods used to estimate LPR here and those in FEAST have several problematic assumptions. Given that NRR is implicitly measured in our LPR estimate, NRR functionality is disabled for this demonstration.

The case study evaluates equivalence of two alt-LDAR programs (P_1 and P_2) against a reference reg-LDAR program (P_{Ref}) configured according to ECCC regulations (Government of Canada, 2018). Three additional programs (P_W, P_L, $P_{W,L}$) modify P_{Ref} to illustrate the importance of environmental envelopes and labour constraints (Table 4.1). The P_W scenario includes an OGI company that limits work in inclement conditions but hires more crews. The P_L scenario represents a company that will operate in a much broader environmental envelope. The $P_{W,L}$ scenario combines P_W and P_L. The ECCC P_{Ref} does not specify environmental envelopes; we assume that most service providers will work in all but the most extreme conditions: temperatures > -30 °C and winds < 20 $m \cdot s^{-1}$. In P_W and $P_{W,L}$ agents are only deployed under preferred weather conditions: > 0 °C, < 5 $m \cdot s^{-1}$ wind, and < 10 mm daily precipitation rate). These thresholds are based loosely on detection modeled performances but are not grounded in legislation or best practices (Ravikumar and Brandt, 2017). Unfortunately, insufficient empirical data exist to inform OGI operational envelopes. Labour constraints are imposed on P_L and $P_{W,L}$ (one OGI crew instead of three). Neither of the screening programs (P_1 and P_2) are constrained by weather or labour. Both P_1 and P_2 use the P_{Ref} OGI method configuration for follow-up. Although more complex programs comprising multiple methods are possible, this demonstration focuses on simple programs.

Close-range methods (used in P_{Ref}, P_W, P_L, $P_{W,L}$, and follow-up) use sigmoidal detection probability curves developed empirically in Ravikumar et al. (2018). For each leak surveyed, shape parameters are sampled from normal distributions $k \sim \mathcal{N}(4.9, 0.3)$ and $x_0 \sim \mathcal{N}(0.47, 0.01)$, assuming an imaging distance of 3 m. We use single-value detection limits for P_1 and P_2 (100 and 5000 $g \cdot h^{-1}$), as empirical detection probability curves have not been published for MGLs or aircraft. These values fall within typical ranges reviewed (Fox et al., 2019a). In P_1 and P_2, follow-up crews visit 80% and 50% of flagged sites, respectively. Note that both alt-LDAR programs are based on hypothetical technologies and work practices and should not be considered representative of typical MGL or aircraft-based programs. Our aversion to using existing systems for this case study is intentional to avoid promoting any one solution provider. Further, most screening technologies have evolving skills and poorly defined work practices.

A constant survey time of 120 minutes per site is used for OGI surveys, following literature (Fox et al., 2019a). The per-site times used for P_1 and P_2 are heuristics and not representative of any real system or work practice. After each visit, a ‘time offsite’ increment is added to the workday, which is broadly representative of all driving or flying, breaks, acquiring permits, and unforeseen circumstances. For methods that must travel overland, we do not explicitly account for between-site distances, which would require modeling of complex and often subjective site selection procedures. Instead, we sample from an empirical distribution of between-site durations acquired during field campaigns in Alberta (unpublished University of Calgary data; available in SI). In the absence of empirical data, P_2 uses a constant 10-minute between-site interval.

7.3 Sensitivity analysis

We conducted 50 000 Monte Carlo simulations, each spanning 1095 days plus a 500-day spin-up period. In each run, programs share a set of global parameters (e.g. LPR, repair delays) sampled from broadly representative probability distributions. Method-specific parameters (e.g. detection limits, survey frequency) are sampled independently. The following index (S) is reported:

$$S = \frac{(O_{high} - O_{low})}{O_{\sigma}}$$

The terms used to calculate S depend on whether the SA evaluates a single program (e.g. Figure 7.5) or the comparison between two programs (Figure 4.3). In both cases, all parameters are randomly sampled for each SA simulation in order to account for cross-variable sensitivity and possible interaction effects. Given that each parameter is evaluated across a broad range of possible scenarios, many of the SA realizations represent unlikely combinations of input parameters, leading to high output variability that can be used to identify sensitive parameters. Here, variability in outputs that affect equivalence determinations are of greatest interest. However, we begin with the definition of S for a single program, which is simpler: O_{high} and O_{low} are the median outputs corresponding to the upper and lower deciles (highest and lowest 10%) of the distribution of input samples. The output standard deviation O_{σ} standardizes S, allowing comparison among different input-output combinations. Put generally, sensitivity of an input parameter is the difference between outputs for the highest inputs and lowest inputs, averaged across all other parameter combinations, and measured in standard deviations of the output distribution. If the difference in outputs is high, so is S, and the input variable has a strong effect.

Calculating S for a SA in which two programs are compared is more complicated. The purpose of equivalence is to compare two LDAR programs. Ideally, the results of that comparison will remain relatively robust to changes in inputs, especially that that must be estimated (e.g., LPR). The guiding principle of the two-program SA is to compare two programs for high input values and compare them again for low input values. Each of these comparisons produces an equivalence metric (e.g., number of leaks or average emission rate). If the metrics are similar for different input values, sensitivity is low. Here, O_{high} and O_{low} are the program comparisons, which are calculated both as a difference and then as the ratio of outputs between two distinct programs sharing the same global parameters. Thus, O should approach zero when programs are equivalent, and S becomes a difference (or ratio) of equivalence measured in standard deviations of the equivalence metric. When S is near zero, the equivalence metric (i.e. difference or ratio of outputs) is independent of the input value.

Two exceptions exist. For binary inputs (e.g. whether venting is considered), median outputs are taken for True/False categories, as deciles are not applicable. When inputs are conditional (e.g. site-level emissions estimates), only inputs that meet the associated condition are considered (e.g. venting is True). Users of LDAR-Sim should note that the SA performed here is specific to a unique set of empirical inputs, environmental conditions, and hypothetical detection methods. We recommend performing a SA each time a new program is evaluated.

Input distributions used for the sensitivity analysis (SA) are shown in Figure 7.6. All input parameters are listed below. Variables with the suffix '_outliers' or '_samples' are used to permute empirical distributions. These distributions include the leak-size distribution (LSD), leak-count distribution (LCD), site-level emission rate distribution (site_rate), and offsite driving times (offsite_times). For each distribution, a number of outliers (integers typically between 0 and 3) is sampled and either added to or removed from the distribution. For example, if -2 was sampled, the two largest elements in the empirical input distribution would be removed. When a new outlier is added, the value is chosen by doubling the largest element in the distribution. For example, if 2 was sampled, two new outliers would be added to the distribution. The second outlier added would be four times the size of the largest element in the original distribution. Input distributions are then sampled from a truncated normal distribution with a mean equal to the length of the original distribution (l) and a standard deviation of one quarter of this value. A lower limit of 10 samples is enforced to ensure proper functioning of the model.

Leak production rate (LPR) is parameterized by fitting a gamma distribution to the empirical distribution of LPR values. For the detection limits listed below, it should be noted that the OGI parameter estimated is variable x_0 described in Ravikumar et al.,(Ravikumar et al., 2018) while the MGL detection limit is an absolute minimum measured in g h^{-1}. For the variables that require integers, samples are rounded as necessary.

$$\text{LSD outliers} \sim Normal(0, 1)$$
$$\text{LSD samples} \sim Normal(l, l/4)$$
$$\text{LCD outliers} \sim Normal(0, 1)$$
$$\text{LCD samples} \sim Normal(l, l/4)$$
$$\text{Site rate outliers} \sim Normal(0,1)$$
$$\text{Site rate samples} \sim Normal(l, l/4)$$
$$\text{Offsite times outliers} \sim Normal(0, 1)$$
$$\text{Offsite times samples} \sim Normal(l, l/4)$$
$$\text{LPR} \sim Gamma(0.8328, 0.03139)$$
$$\text{Repair delay} \sim Uniform(0, 100)$$

$$\text{Operator strength } \sim Exponential(0.1)$$

$$\text{Operator bonus } \sim Exponential(0.2)$$

$$\text{Consider operator } \sim Binomial(1, 0.2)$$

$$\text{Consider daylight } \sim Binomial(1, 0.5)$$

$$\text{Consider venting } \sim Binomial(1, 0.5)$$

$$\text{Maximum work day } \sim Uniform(6, 14)$$

$$\text{OGI number of crews } \sim Poisson(0.5) + 1$$

$$\text{OGI minimum temperature } \sim Normal(-20, 10)$$

$$\text{OGI maximum wind } \sim Normal(15, 3)$$

$$\text{OGI maximum precipitation } \sim Uniform(0, 0.1)$$

$$\text{OGI reporting delay } \sim Uniform(0, 30)$$

$$\text{OGI time per site } \sim Uniform(30, 500)$$

$$\text{OGI required surveys } \sim Uniform(1, 4)$$

$$\text{OGI minimum survey interval } \sim Uniform(0, 90)$$

$$\text{OGI detection limit } \sim Uniform(0, 2)$$

$$\text{MGL number of crews } \sim Poisson(0.5) + 1$$

$$\text{MGL minimum temperature } \sim Normal(-20, 10)$$

$$\text{MGL maximum wind } \sim Normal(15, 3)$$

$$\text{MGL maximum precipitation } \sim Uniform(0, 0.1)$$

$$\text{MGL reporting delay } \sim Uniform(0, 30)$$

$$\text{MGL time per site } \sim Uniform(1,30)$$

$$\text{MGL required surveys } \sim Uniform(1, 4)$$

$$\text{MGL minimum survey interval } \sim Uniform(0, 90)$$

$$\text{MGL detection limit } \sim Uniform(1, 100)$$

$$\text{MGL followup threshold } \sim Uniform(0, 500)$$

$$\text{MGL followup ratio } \sim Uniform(0.1, 1)$$

7.4 Supporting tables and figures

Table 7.1 LDAR-Sim output data by day, site, and leak.

By day	By site	By leak
date	identifying information*	identifying information*
emissions	emissions	emission rate
new leak count	initial leak count	date began
active leak count	active leak count	date found
repaired leak count	repaired leak count	date repaired
tag count	flag status	days active
proportion sites available†	flag date	repair delay
cost†	flag company	found by (company)
redundant tags‡	missed leaks†	found by (crew)
effective flags§	surveys conducted†	status‖
redundant flags 1§	days since last survey†	
redundant flags 2§		
redundant flags 3§		

* Each site and leak are assigned a unique identifier. Site name, operator, and location are output if provided by the user in the input file.

† Reported by method

‡ Reported only for close-range methods and operator

§ Reported only for screening methods

‖ Can be active, repaired, or tagged

Table 7.2 Case study results

	P_{Ref}	P_W	P_L	$P_{W,\,L}$	P_1	P_2
Method	OGI	OGI	OGI	OGI	Truck	Aircraft
Proportion sites available	1.0 ± 0.1	0.4 ± 0.4	1.0 ± 0.1	0.4 ± 0.4	1.0 ± 0.1	1.0 ± 0.1
Sites flagged per day	-	-	-	-	7.3 ± 1.2	2.6 ± 2.6
Leaks tagged per day	7.5 ± 4.9	7.6 ± 8.2	7.6 ± 3.0	7.6 ± 6.4	7.5 ± 3.0	7.0 ± 7.4
Emissions (kg · site^{-1} · day^{-1})	4.0 ± 1.2	5.9 ± 2.3	8.4 ± 1.5	12.3 ± 2.7	4.4 ± 1.1	11.9 ± 1.6

Table 7.3 Approximate number of active leaks and daily emissions for 500 sites under a range of p_{max} values (simulations conducted without LDAR). Note that if the largest leaks have only a 5% chance of being detected by operators during weekly visits (gold line in Figure 7.3), system-wide emissions fall by >50%.

	0.00	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**
Active leaks	3400	3100	2800	2500	2400	2400
Emissions (kg)	49000	20000	13000	9000	8000	7000

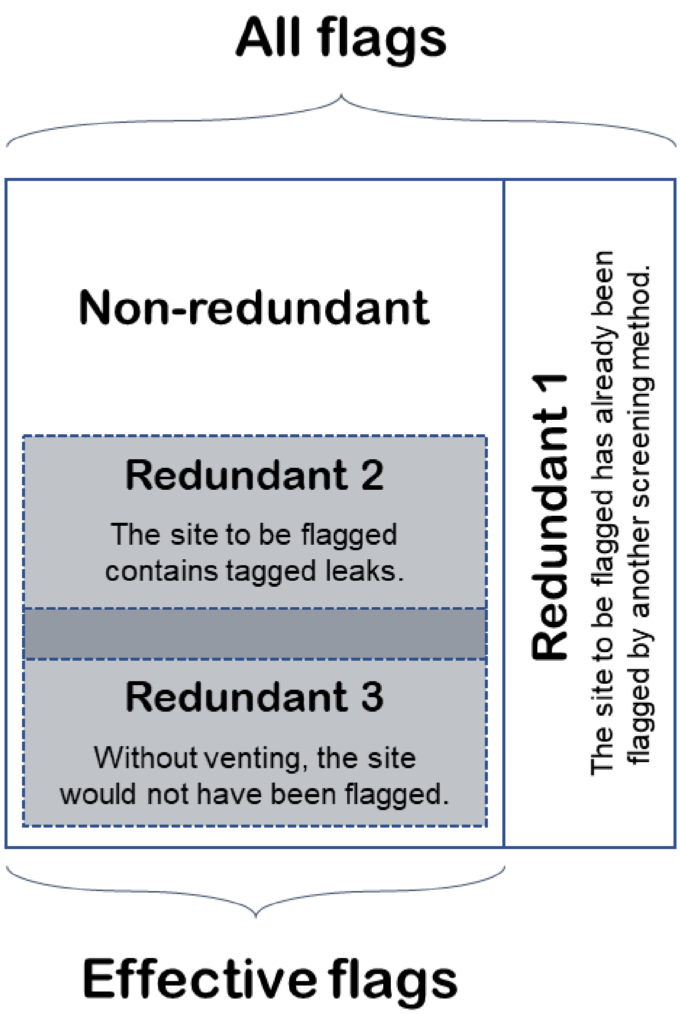

Figure 7.1 Illustration of the three types of redundancies tracked in LDAR-Sim.

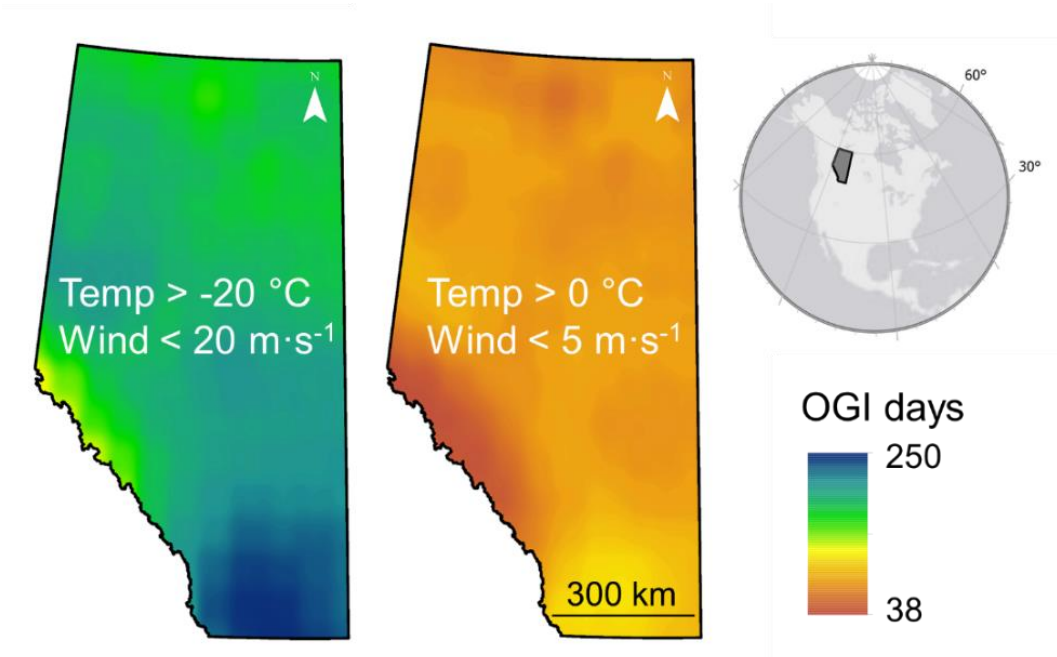

Figure 7.2 The spatial distribution of OGI suitability in Alberta for different deployment thresholds. On the left, crews are deployed if warmer than -20 °C and winds are below 20 $m \cdot s^{-1}$. On the right, stricter thresholds are used, resulting in fewer deployment days.

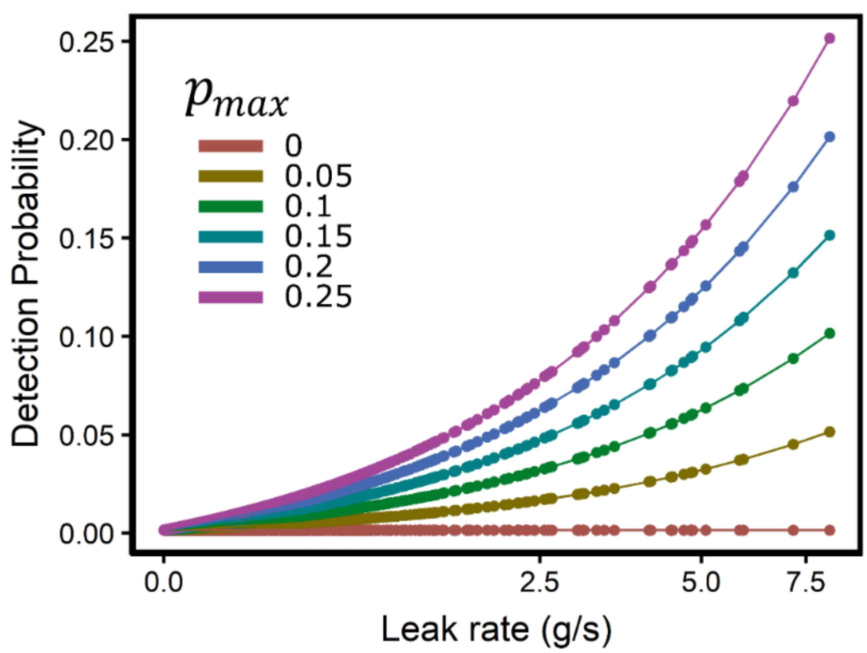

Figure 7.3 Detection probability curves for six possible $\boldsymbol{p_{max}}$ values using the Fort Worth Air Quality Study. For example, the gold line ($\boldsymbol{p_{max}}$ = 0.05) assumes operators have a 5% chance of detecting the largest leak in the dataset in NRR calculations.

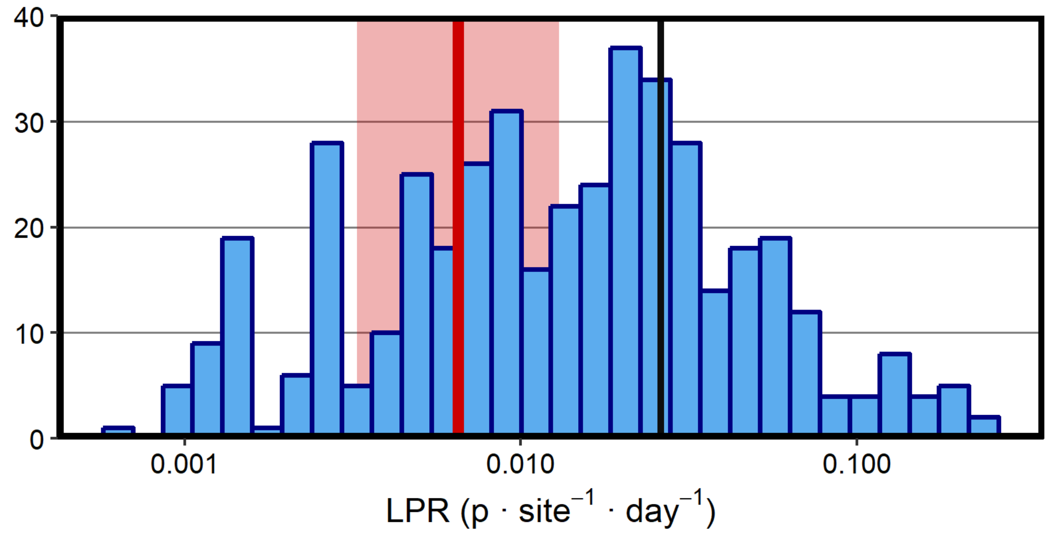

Figure 7.4 Empirical frequency distribution of leak production rates in Alberta. The mean (vertical black line) is 0.026. The red line is the LPR value used in FEAST and in this LDAR-Sim demonstration. The shaded region is the LPR sensitivity analysis range evaluated in Kemp et al. (2016).

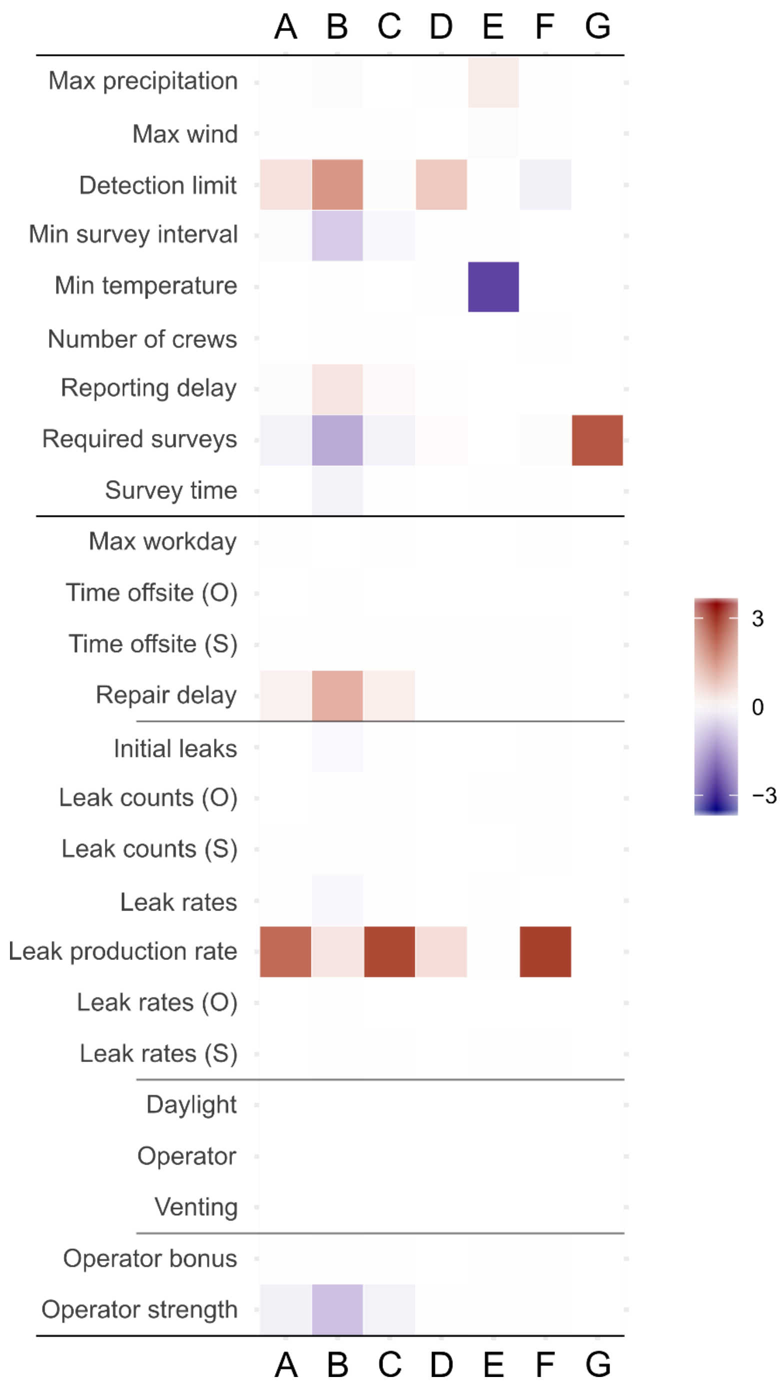

Figure 7.5 SA results for a single OGI program. Rows are inputs and columns are outputs: (A) number of active leaks, (B) number of days leaks are active, (C) daily emissions, (D) missed leaks, (E) proportion of sites available for LDAR, (F) repaired leaks, and (G) surveys conducted.

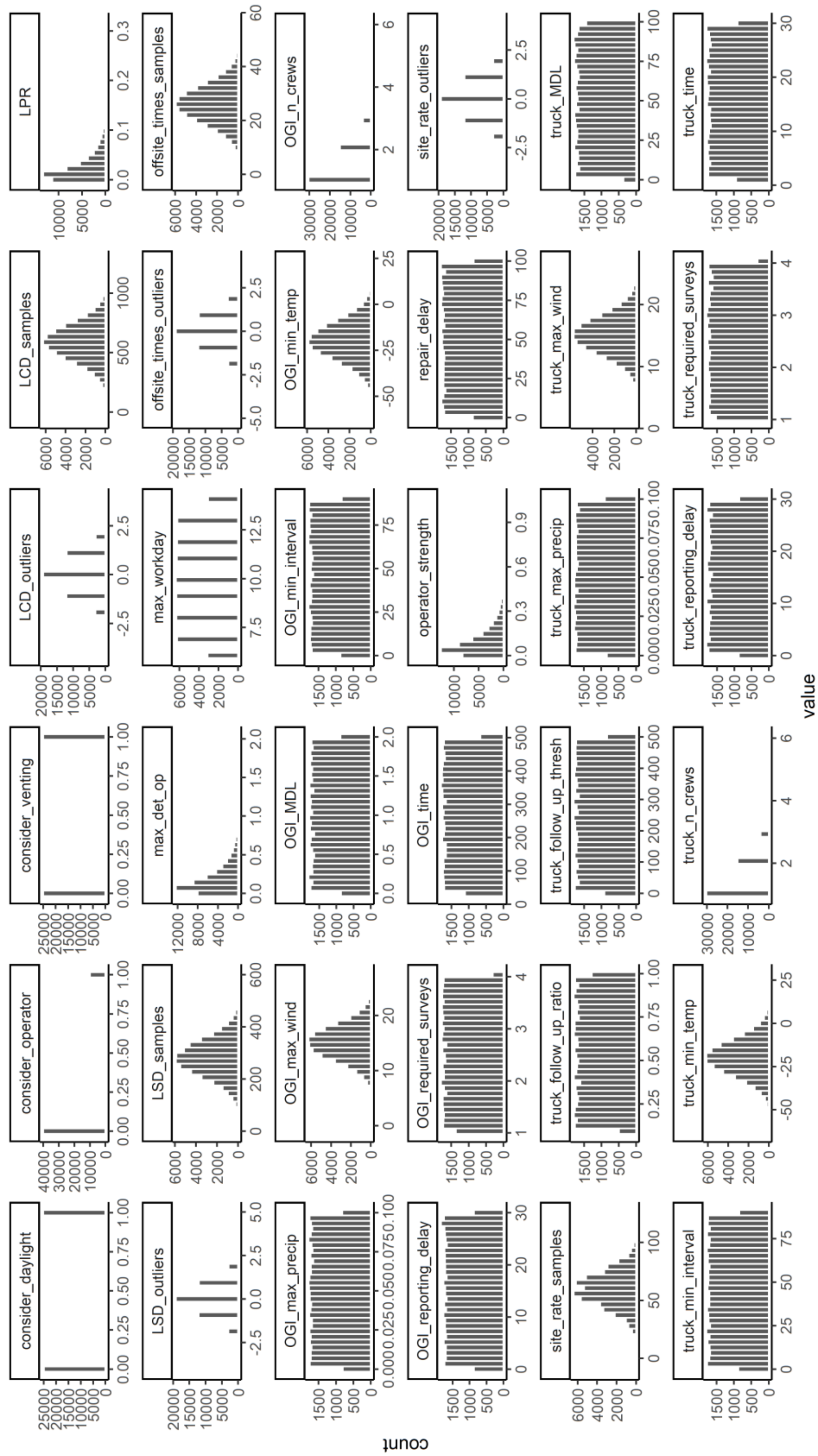

Figure 7.6 Sampling distributions used in the SA (N = 50 000).

Chapter 8

Appendix B: Supplementary information for Chapter 5

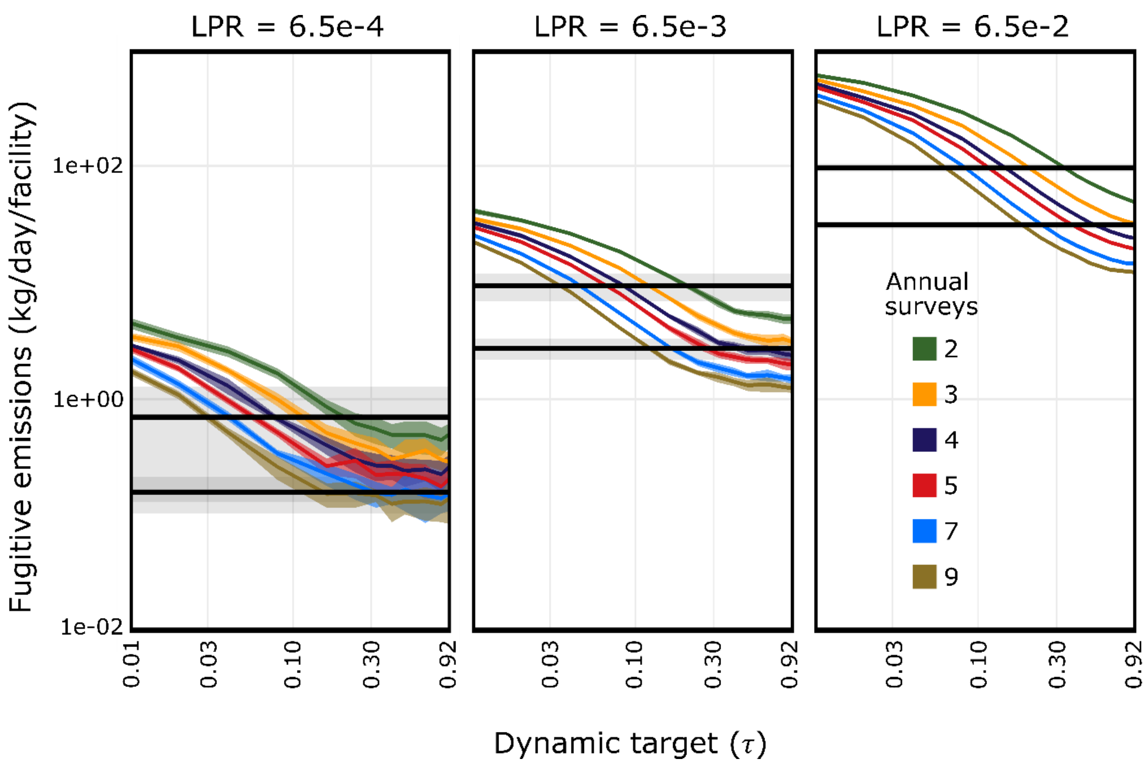

Figure 8.1 Equivalence scenarios (D1; E = 1; no design emissions) for leak production rates one order of magnitude below (0.00065, left) and above (0.065, right) the LPR used for this study (0.0065, middle). All simulations performed without null repair.

Chapter 9

Literature cited

AER, 2020. ST37: List of Wells in Alberta. Alberta Energy Regulator.

AER, 2019. ST98: Alberta Energy Outlook. Alberta Energy Regulator.

AER, 2018. Directive 060: Upstream petroleum industry flaring, incinerating, and venting. Alberta Energy Regulator.

Albertson, John.D., Harvey, T., Foderaro, G., Zhu, P., Zhou, X., Ferrari, S., Amin, M.S., Modrak, M., Brantley, H., Thoma, E.D., 2016. A Mobile Sensing Approach for Regional Surveillance of Fugitive Methane Emissions in Oil and Gas Production. Environ. Sci. Technol. 50, 2487–2497. https://doi.org/10.1021/acs.est.5b05059

Alden, C.B., Ghosh, S., Coburn, S., Sweeney, C., Karion, A., Wright, R., Coddington, I., Prasad, K., Rieker, G.B., 2017. Methane leak detection and sizing over long distances using dual frequency comb laser spectroscopy and a bootstrap inversion. Atmospheric Meas. Tech. Discuss.

Allen, D.T., Cardoso Saldaña, F.J., Kimura, Y., 2017. Variability in Spatially and Temporally Resolved Emissions and Hydrocarbon Source Fingerprints for Oil and Gas Sources in Shale Gas Production Regions. Environ. Sci. Technol.

Allen, D.T., Pacsi, A.P., Sullivan, D.W., Zavala-Araiza, D., Harrison, M., Keen, K., Fraser, M.P., Daniel Hill, A., Sawyer, R.F., Seinfeld, J.H., 2014a. Methane emissions from process equipment at natural gas production sites in the United States: Pneumatic controllers. Environ. Sci. Technol. 49, 633–640.

Allen, D.T., Sullivan, D.W., Zavala-Araiza, D., Pacsi, A.P., Harrison, M., Keen, K., Fraser, M.P., Daniel Hill, A., Lamb, B.K., Sawyer, R.F., others, 2014b. Methane emissions from process equipment at natural gas production sites in the United States: Liquid unloadings. Environ. Sci. Technol. 49, 641–648.

Allen, D.T., Torres, V.M., Thomas, J., Sullivan, D.W., Harrison, M., Hendler, A., Herndon, S.C., Kolb, C.E., Fraser, M.P., Hill, A.D., others, 2013. Measurements of methane emissions at natural gas production sites in the United States. Proc. Natl. Acad. Sci. 110, 17768–17773.

Alvarez, R.A., Zavala-Araiza, D., Lyon, D.R., Allen, D.T., Barkley, Z.R., Brandt, A.R., Davis, K.J., Herndon, S.C., Jacob, D.J., Karion, A., 2018. Assessment of methane emissions from the US oil and gas supply chain. Science 361, 186–188. https://doi.org/10.1126/science.aar7204

ARPA-E, 2018. MONITOR Projects [WWW Document]. URL https://arpa-e.energy.gov/?q=program-projects/MONITOR (accessed 10.21.18).

Atherton, E., Risk, D., Fougere, C., Lavoie, M., Marshall, A., Werring, J., Williams, J.P., Minions, C., 2017. Mobile measurement of methane emissions from natural gas developments in Northeastern British Columbia, Canada. Atmospheric Chem. Phys. Discuss. 1–28. https://doi.org/10.5194/acp-2017-109

Babilotte, A., Lagier, T., Fiani, E., Taramini, V., 2010. Fugitive methane emissions from landfills: Field comparison of five methods on a French landfill. J. Environ. Eng. 136, 777–784.

Barchyn, T., Hugenholtz, C.H., Fox, T.A., 2019. Plume detection modeling of a drone-based natural gas leak detection system. Elem Sci Anth 7.

Barchyn, T., Hugenholtz, C.H., Myshak, S., Bauer, J., 2017. A UAV-based system for detecting natural gas leaks. J. Unmanned Veh. Syst. 6, 18–30.

Bartholomew, J., Lyman, P., Weimer, C., Tandy, W., 2017. Wide area methane emissions mapping with airborne IPDA lidar, in: Lidar Remote Sensing for Environmental Monitoring 2017. International Society for Optics and Photonics, p. 1040607.

Bateman, J., Zekkos, D., Olson, E., Messenger, S.M., Kershaw, C., Fei, X., Lynch, J., 2016. Preliminary Observations from Robot-Enabled Surface Methane Concentration Monitoring at a MSW Landfill, in: Geo-Chicago 2016. pp. 740–749.

Bell, C., Vaughn, T., Zimmerle, D., Herndon, S., Yacovitch, T., Heath, G., Pétron, G., Edie, R., Field, R., Murphy, S., 2017. Comparison of methane emission estimates from multiple measurement techniques at natural gas production pads. Elem Sci Anth 5.

Benson, R., Madding, R., Lucier, R., Lyons, J., Czerepuszko, P., 2006. Standoff passive optical leak detection of volatile organic compounds using a cooled InSb based infrared imager, in: AWMA 99th Annual Meeting Papers.

Boucher, O., Friedlingstein, P., Collins, B., Shine, K.P., 2009. The indirect global warming potential and global temperature change potential due to methane oxidation. Environ. Res. Lett. 4, 044007.

Brandt, A.R., Heath, G.A., Cooley, D., 2016. Methane leaks from natural gas systems follow extreme distributions. Environ. Sci. Technol. 50, 12512–12520.

Brandt, A.R., Heath, G.A., Kort, E.A., O'Sullivan, F., Pétron, G., Jordaan, S.M., Tans, P., Wilcox, J., Gopstein, A.M., Arent, D., others, 2014. Methane leaks from North American natural gas systems. Science 343, 733–735.

Brandt, A.R., Pétron, G., 2015. Fugitive emissions and air quality impacts of US natural gas systems. The Bridge 45, 22–31.

Brantley, H.L., Thoma, E.D., Squier, W.C., Guven, B.B., Lyon, D., 2014. Assessment of methane emissions from oil and gas production pads using mobile measurements. Environ. Sci. Technol. 48, 14508–14515.

Brownlow, R., Lowry, D., Thomas, R.M., Fisher, R.E., France, J.L., Cain, M., Richardson, T.S., Greatwood, C., Freer, J., Pyle, J.A., MacKenzie, A.R., Nisbet, E.G., 2016. Methane mole fraction and $\delta^{13}C$ above and below the trade wind inversion at Ascension Island in air sampled by aerial robotics. Geophys. Res. Lett. 43, 11,893-11,902. https://doi.org/10.1002/2016GL071155

Bruhwiler, L.M., Basu, S., Bergamaschi, P., Bousquet, P., Dlugokencky, E., Houweling, S., Ishizawa, M., Kim, H.-S., Locatelli, R., Maksyutov, S., 2017. US CH4 emissions from oil and gas production: Have recent large increases been detected? J. Geophys. Res. Atmospheres 122, 4070–4083.

Buckland, K.N., Young, S.J., Keim, E.R., Johnson, B.R., Johnson, P.D., Tratt, D.M., 2017. Tracking and quantification of gaseous chemical plumes from anthropogenic emission sources within the Los Angeles Basin. Remote Sens. Environ. 201, 275–296.

Buckland, S.T., Burnham, K.P., Augustin, N.H., 1997. Model selection: an integral part of inference. Biometrics 603–618.

Butler, T.M., Simmonds, I., Rayner, P.J., 2004. Mass balance inverse modelling of methane in the 1990s using a Chemistry Transport Model. Atmospheric Chem. Phys. 4, 2561–2580.

Butz, A., Galli, A., Hasekamp, O., Landgraf, J., Tol, P., Aben, I., 2012. TROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH4 retrievals for aerosol and cirrus loaded atmospheres. Remote Sens. Environ. 120, 267–276.

CAPP, 2007. Management of fugitive emissions at upstream oil and gas facilities. Canadian Association of Petroleum Producers.

CATF, 2013. Methane emissions from the oil and gas sector: Barriers to abatement and technologies for emission reductions. Clean Air Task Force, Boston, MA.

Caulton, D.R., Li, Q., Bou-Zeid, E., Fitts, J.P., Golston, L.M., Pan, D., Lu, J., Lane, H.M., Buchholz, B., Guo, X., 2018. Quantifying uncertainties from mobile-laboratory-derived emissions of well pads using inverse Gaussian methods. Atmospheric Chem. Phys. 18, 15145–15168.

CDPHE, 2018. Alternative AIMM Guidance & Procedures.

Christen, A., 2014. Atmospheric measurement techniques to quantify greenhouse gas emissions from cities. Urban Clim. 10, 241–260.

Clearstone Engineering Ltd., 2018. Update of Equipment, Component and Fugitive Emission Factors for Alberta Upstream Oil and Gas. (Technical Report).

Coburn, S., Alden, C.B., Wright, R., Cossel, K., Baumann, E., Truong, G.-W., Giorgetta, F., Sweeney, C., Newbury, N.R., Prasad, K., 2017. Continuous regional trace gas source attribution using a field-deployed dual frequency comb spectrometer. ArXiv Prepr. ArXiv171108067.

Coburn, S., Alden, C.B., Wright, R., Cossel, K., Baumann, E., Truong, G.-W., Giorgetta, F., Sweeney, C., Newbury, N.R., Prasad, K., Coddington, I., Rieker, G.B., 2018. Regional trace-gas source attribution using a field-deployed dual frequency comb spectrometer. Optica 5, 320–327. https://doi.org/10.1364/OPTICA.5.000320

Conley, S., Faloona, I., Mehrotra, S., Suard, M., Lenschow, D.H., Sweeney, C., Herndon, S., Schwietzke, S., Pétron, G., Pifer, J., Kort, E.A., Schnell, R., 2017. Application of Gauss's Theorem to quantify localized surface emissions from airborne measurements of wind and trace gases. Atmospheric Meas. Tech. Discuss. 1–29. https://doi.org/10.5194/amt-2017-55

Conley, S., Franco, G., Faloona, I., Blake, D.R., Peischl, J., Ryerson, T.B., 2016. Methane emissions from the 2015 Aliso Canyon blowout in Los Angeles, CA. Science 351, 1317–1320.

Cossel, K.C., Waxman, E.M., Giorgetta, F.R., Cermak, M., Coddington, I.R., Hesselius, D., Ruben, S., Swann, W.C., Truong, G.-W., Rieker, G.B., Newbury, N.R., 2017. Open-path dual-comb spectroscopy to an airborne retroreflector. Optica 4, 724–728. https://doi.org/10.1364/OPTICA.4.000724

Davis, K.J., Deng, A., Lauvaux, T., Miles, N.L., Richardson, S.J., Sarmiento, D.P., Gurney, K.R., Hardesty, R.M., Bonin, T.A., Brewer, W.A., Lamb, B.K., Shepson, P.B., Harvey, R.M., Cambaliza, M.O., Sweeney, C., Turnbull, J.C., Whetstone, J., Karion, A., 2017. The Indianapolis Flux Experiment (INFLUX): A test-bed for developing urban greenhouse gas emission measurements. Elem Sci Anth 5, 21. https://doi.org/10.1525/elementa.188

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., Berg, L. van de, Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., Rosnay, P. de, Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137, 553–597. https://doi.org/10.1002/qj.828

Duren, R.M., Thorpe, A.K., Foster, K.T., Rafiq, T., Hopkins, F.M., Yadav, V., Bue, B.D., Thompson, D.R., Conley, S., Colombi, N.K., 2019. California's methane super-emitters. Nature 575, 180–184.

ECCC, 2018. SOR/2018-66: Regulations Respecting Reduction in the Release of Methane and Certain Volatile Organic Compounds (Upstream Oil and Gas Sector).

ECCC, 2016. National Inventory Report 1990−2014: Greenhouse gas sources and sinks in Canada Part 3. Environment and Climate Change Canada, Ottawa.

Ellis, H.M., Lackaye, R., 1989. Estimating fugitive emissions of volatile compounds from equipment leaks. JAPCA 39, 1619–1622.

Emran, B.J., Tannant, D.D., Najjaran, H., 2017. Low-Altitude Aerial Methane Concentration Mapping. Remote Sens. 9, 823.

Englander, J.G., Brandt, A.R., Conley, S., Lyon, D.R., Jackson, R.B., 2018. Aerial Interyear Comparison and Quantification of Methane Emissions Persistence in the Bakken Formation of North Dakota, USA. Environ. Sci. Technol. 52, 8947–8953.

Environmental Defense Fund, 2016. Finding, fixing leaks is a cost-effectiveway to cut oil and gas methane emissions.

Envirotech Engineering, 2007. Review and update of methods used for air emissions leak detection and quantification. Petroleum Technology Alliance Canada.

EPA, 2019. Inventory of U.S. Greenhouse Gas Emissions and Sinks, 1990-2017. US Environmental Protection Agency.

EPA, 2016. Oil and natural gas sector: Emission standards for new, reconstructed, and modified sources (Federal register: Final rule). Government of the United States of America.

Etheridge, D.M., Steele, L.P., Francey, R.J., Langenfelds, R.L., 1998. Atmospheric methane between 1000 A.D. and present: Evidence of anthropogenic emissions and climatic variability. J. Geophys. Res. Atmospheres 103, 15979–15993. https://doi.org/10.1029/98JD00923

Feitz, A., Schroder, I., Phillips, F., Coates, T., Neghandhi, K., Day, S., Luhar, A., Bhatia, S., Edwards, G., Hrabar, S., 2018. The Ginninderra CH 4 and CO 2 release experiment: An evaluation of gas detection and quantification techniques. Int. J. Greenh. Gas Control 70, 202–224.

Footer, T.L., 2015. Optical gas imaging protocol (Technical support document). Eastern Research Group, Inc.

Foster-Wittig, T., 2015. Mobile Sensors: Assessment of Fugitive Methane Emissions from Near and Far-Field Sources. Duke University.

Fox, T.A., Barchyn, T.E., Risk, D., Ravikumar, A.P., Hugenholtz, C.H., 2019a. A review of close-range and screening technologies for mitigating fugitive methane emissions in upstream oil and gas. Environ. Res. Lett. 14, 053002.

Fox, T.A., Gao, M., Barchyn, T.E., Jamin, Y., Hugenholtz, C.H., in review. An agent-based model for estimating emissions reduction equivalence among leak detection and repair programs.

Fox, T.A., Ravikumar, A.P., Hugenholtz, C.H., Zimmerle, D., Barchyn, T.E., Johnson, M., Lyon, D., Taylor, T., 2019b. A methane emissions reduction equivalence framework for alternative leak detection and repair programs. Elementa 7.

Frankenberg, C., Aben, I., Bergamaschi, P., Dlugokencky, E.J., Van Hees, R., Houweling, S., Van Der Meer, P., Snel, R., Tol, P., 2011. Global column-averaged methane mixing ratios from 2003 to 2009 as derived from SCIAMACHY: Trends and variability. J. Geophys. Res. Atmospheres 116.

Frankenberg, C., Meirink, J.F., van Weele, M., Platt, U., Wagner, T., 2005. Assessing methane emissions from global space-borne observations. Science 308, 1010–1014.

Frankenberg, C., Thorpe, A.K., Thompson, D.R., Hulley, G., Kort, E.A., Vance, N., Borchardt, J., Krings, T., Gerilowski, K., Sweeney, C., Conley, S., Bue, B.D., Aubrey, A.D., Hook, S., Green, R.O., 2016. Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region. Proc. Natl. Acad. Sci. 113, 9734–9739. https://doi.org/10.1073/pnas.1605617113

Gålfalk, M., Olofsson, G., Crill, P., Bastviken, D., 2016. Making methane visible. Nat. Clim. Change 6, 426–430. https://doi.org/10.1038/nclimate2877

Garcia-Gonzales, D.A., Shonkoff, S.B., Hays, J., Jerrett, M., 2019. Hazardous air pollutants associated with upstream oil and natural gas development: A critical synthesis of current peer-reviewed literature. Annu. Rev. Public Health 40, 283–304.

Gardi, A., Kapoor, R., Sabatini, R., 2017. Detection of Volatile Organic Compound Emissions from Energy Distribution Network Leaks by Bistatic LIDAR. Energy Procedia 110, 396–401. https://doi.org/10.1016/j.egypro.2017.03.159

Gemerek, J.R., Ferrari, S., Albertson, J.D., 2017. Fugitive gas emission rate estimation using multiple heterogeneous mobile sensors, in: Olfaction and Electronic Nose (ISOEN), 2017 ISOCS/IEEE International Symposium On. IEEE, pp. 1–3.

Gerhart, T., Sunu, J., Lieu, L., Merkurjev, E., Chang, J.-M., Gilles, J., Bertozzi, A.L., 2013. Detection and tracking of gas plumes in LWIR hyperspectral video sequence data. SPIE Def. Secur. Sens. 87430J–87430J.

Goetz, J.D., Avery, A., Werden, B., Floerchinger, C., Fortner, E.C., Wormhoudt, J., Massoli, P., Herndon, S.C., Kolb, C.E., Knighton, W.B., Peischl, J., Warneke, C., De Gouw, J.A., Shaw, S.L., DeCarlo, P.F., 2017. Analysis of local-scale background concentrations of methane and other gas-phase species in the Marcellus Shale. Elem Sci Anth 5, 1. https://doi.org/10.1525/elementa.182

Goldsmith, C.D., Chanton, J., Abichou, T., Swan, N., Green, R., Hater, G., 2012. Methane emissions from 20 landfills across the United States using vertical radial plume mapping. J. Air Waste Manag. Assoc. 62, 183–197. https://doi.org/10.1080/10473289.2011.639480

Golston, L.M., Aubut, N., Frish, M., Yang, S., Talbot, R., Gretencord, C., McSpiritt, J., Zondlo, M., 2018. Natural Gas Fugitive Leak Detection Using an Unmanned Aerial Vehicle: Localization and Quantification of Emission Rate. Atmosphere 9, 333.

Golston, L.M., Tao, L., Brosy, C., Sch?fer, K., Wolf, B., McSpiritt, J., Buchholz, B., Caulton, D.R., Pan, D., Zondlo, M.A., Yoel, D., Kunstmann, H., McGregor, M., 2017. Lightweight mid-infrared methane sensor for unmanned aerial systems. Appl. Phys. B 123. https://doi.org/10.1007/s00340-017-6735-6

Government of Canada, 2018. Regulations Respecting Reduction in the Release of Methane and Certain Volatile Organic Compounds (Upstream Oil and Gas Sector) (SOR/2018-66).

Government of Canada, E. and C.C.C., 2016. Environment and Climate Change Canada - Climate Change - National Inventory Report 1990-2014: Greenhouse Gas Sources and Sinks in Canada - Executive Summary [WWW Document]. URL https://www.ec.gc.ca/ges-ghg/default.asp?lang=En&n=662F9C56-1 (accessed 7.5.17).

GreenPath Energy, 2017. Historical Canadian Fugitive Emissions Management Program Assessment (Technical Report). Petroleum Technology Alliance Canada.

Harriss, R., Alvarez, R.A., Lyon, D., Zavala-Araiza, D., Nelson, D., Hamburg, S.P., 2015. Using Multi-Scale Measurements to Improve Methane Emission Estimates from Oil and Gas Operations in the Barnett Shale Region, Texas. Environ. Sci. Technol. 49, 7524–7526. https://doi.org/10.1021/acs.est.5b02305

Hashmonay, R.A., Yost, M.G., 1999. Innovative Approach for Estimating Fugitive Gaseous Fluxes Using Computed Tomography and Remote Optical Sensing Techniques. J. Air Waste Manag. Assoc. 49, 966–972. https://doi.org/10.1080/10473289.1999.10463862

Heimburger, A.M., Harvey, R.M., Shepson, P.B., Stirm, B.H., Gore, C., Turnbull, J., Cambaliza, M.O., Salmon, O.E., Kerlo, A.-E.M., Lavoie, T.N., others, 2017. Assessing the optimized precision of the aircraft mass balance method for measurement of urban greenhouse gas emission rates through averaging. Elem Sci Anth 5.

Hirst, B., Jonathan, P., del Cueto, F.G., Randell, D., Kosut, O., 2013. Locating and quantifying gas emission sources using remotely obtained concentration data. Atmos. Environ. 74, 141–158.

Hirst, B., Randell, D., Jones, M., Jonathan, P., King, B., Dean, M., 2017. A new technique for monitoring the atmosphere above onshore carbon storage projects that can estimate the locations and mass emission rates of detected sources. Energy Procedia 114, 3716–3728.

Hu, H., Landgraf, J., Detmers, R., Borsdorff, T., Aan de Brugh, J., Aben, I., Butz, A., Hasekamp, O., 2018. Toward global mapping of methane with TROPOMI: first results and intersatellite comparison to GOSAT. Geophys. Res. Lett. 45, 3682–3689.

ICF International, 2015. Economic Analysis of Methane Emission Reduction Opportunities in the Canadian Oil and Natural Gas Industries (Economic Analysis). Environmental Defense Fund.

ICF International, 2014. Economic Analysis of Methane Emission Reduction Opportunities in the U.S. Onshore Oil and Natural Gas Industries (Economic Analysis). Environmental Defense Fund.

International Energy Agency, 2017. World Energy Outlook.

IPCC, 2018. Global Warming of 1.5 ºC (Special Report). The Intergovernmental Panel on Climate Change.

Jackson, R.B., Down, A., Phillips, N.G., Ackley, R.C., Cook, C.W., Plata, D.L., Zhao, K., 2014. Natural Gas Pipeline Leaks Across Washington, DC. Environ. Sci. Technol. 48, 2051–2058. https://doi.org/10.1021/es404474x

Jacob, D.J., Turner, A.J., Maasakkers, J.D., Sheng, J., Sun, K., Liu, X., Chance, K., Aben, I., McKeever, J., Frankenberg, C., 2016. Satellite observations of atmospheric methane and their value for quantifying methane emissions. Atmospheric Chem. Phys. 16, 14371–14396. https://doi.org/10.5194/acp-16-14371-2016

Johnson, D., Heltzel, R., Oliver, D., 2019. Temporal Variations in Methane Emissions from an Unconventional Well Site. ACS Omega 4, 3708–3715.

Johnson, D.R., Covington, A.N., Clark, N.N., 2015. Methane Emissions from Leak and Loss Audits of Natural Gas Compressor Stations and Storage Facilities. Environ. Sci. Technol. 49, 8132–8138. https://doi.org/10.1021/es506163m

Johnson, M.R., Tyner, D.R., 2020. A case study in competing methane regulations: Will Canada's and Alberta's contrasting regulations achieve equivalent reductions? Elem Sci Anth 8.

Johnson, M.R., Tyner, D.R., Conley, S., Schwietzke, S., Zavala-Araiza, D., 2017. Comparisons of Airborne Measurements and Inventory Estimates of Methane Emissions in the Alberta Upstream Oil and Gas Sector. Environ. Sci. Technol.

Kang, M., Kanno, C.M., Reid, M.C., Zhang, X., Mauzerall, D.L., Celia, M.A., Chen, Y., Onstott, T.C., 2014. Direct measurements of methane emissions from abandoned oil and gas wells in Pennsylvania. Proc. Natl. Acad. Sci. 111, 18173–18177.

Karion, A., Sweeney, C., Kort, E.A., Shepson, P.B., Brewer, A., Cambaliza, M., Conley, S.A., Davis, K., Deng, A., Hardesty, M., Herndon, S.C., Lauvaux, T., Lavoie, T., Lyon, D., Newberger, T., P?tron, G., Rella, C., Smith, M., Wolter, S., Yacovitch, T.I., Tans, P., 2015. Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region. Environ. Sci. Technol. 49, 8124–8131. https://doi.org/10.1021/acs.est.5b00217

Karion, A., Sweeney, C., Pétron, G., Frost, G., Michael Hardesty, R., Kofler, J., Miller, B.R., Newberger, T., Wolter, S., Banta, R., others, 2013. Methane emissions estimate from airborne measurements over a western United States natural gas field. Geophys. Res. Lett. 40, 4393–4397.

Kemp, C.E., Ravikumar, A.P., Brandt, A.R., 2016. Comparing Natural Gas Leakage Detection Technologies Using an Open-Source Virtual Gas Field Simulator. Environ. Sci. Technol. 50, 4546–4553. https://doi.org/10.1021/acs.est.5b06068

Khan, A., Schaefer, D., Tao, L., Miller, D.J., Sun, K., Zondlo, M.A., Harrison, W.A., Roscoe, B., Lary, D.J., 2012. Low power greenhouse gas sensors for unmanned aerial vehicles. Remote Sens. 4, 1355–1368.

Kiemle, C., Kawa, S.R., Quatrevalet, M., Browell, E.V., 2014. Performance simulations for a spaceborne methane lidar mission. J. Geophys. Res. Atmospheres 119, 4365–4379.

Kort, E.A., Frankenberg, C., Costigan, K.R., Lindenmaier, R., Dubey, M.K., Wunch, D., 2014. Four corners: The largest US methane anomaly viewed from space: Four Corners: largest US methane anomaly. Geophys. Res. Lett. 41, 6898–6903. https://doi.org/10.1002/2014GL061503

Kuze, A., Shiomi, K., Tanaka, M., Ueda, Y., Deguchi, A., Yoshida, J., Yamamoto, Y., Kataoka, F., Taylor, T.E., Buijs, H., others, 2016. Update on GOSAT TANSO-FTS performance, operations, and data products after more than six years in space. Atmospheric Meas. Tech. 2016, 1–1.

Lan, X., Talbot, R., Laine, P., Torres, A., 2015. Characterizing fugitive methane emissions in the Barnett Shale area using a mobile laboratory. Environ. Sci. Technol. 49, 8139–8146.

Lavoie, T.N., Shepson, P.B., Cambaliza, M.O., Stirm, B.H., Conley, S., Mehrotra, S., Faloona, I.C., Lyon, D., 2017. Spatiotemporal Variability of Methane Emissions at Oil and Natural Gas Operations in the Eagle Ford Basin. Environ. Sci. Technol.

Lavoie, T.N., Shepson, P.B., Cambaliza, M.O., Stirm, B.H., Karion, A., Sweeney, C., Yacovitch, T.I., Herndon, S.C., Lan, X., Lyon, D., 2015. Aircraft-based measurements of point source methane emissions in the Barnett Shale basin. Environ. Sci. Technol. 49, 7904–7913.

Littlefield, J.A., Marriott, J., Schivley, G.A., Skone, T.J., 2017. Synthesis of recent ground-level methane emission measurements from the U.S. natural gas supply chain. J. Clean. Prod. 148, 118–126. https://doi.org/10.1016/j.jclepro.2017.01.101

Lyman, S.N., Tran, T., Mansfield, M.L., Ravikumar, A.P., 2019. Aerial and ground-based optical gas imaging survey of Uinta Basin oil and gas wells. Elem Sci Anth 7.

Lyon, D.R., Alvarez, R.A., Zavala-Araiza, D., Brandt, A.R., Jackson, R.B., Hamburg, S.P., 2016. Aerial surveys of elevated hydrocarbon emissions from oil and gas production sites. Environ. Sci. Technol. 50, 4877–4886.

Malaver, A., Motta, N., Corke, P., Gonzalez, F., 2015. Development and Integration of a Solar Powered Unmanned Aerial Vehicle and a Wireless Sensor Network to Monitor Greenhouse Gases. Sensors 15, 4072–4096. https://doi.org/10.3390/s150204072

Marchese, A.J., Vaughn, T.L., Zimmerle, D.J., Martinez, D.M., Williams, L.L., Robinson, A.L., Mitchell, A.L., Subramanian, R., Tkacik, D.S., Roscioli, J.R., Herndon, S.C., 2015. Methane Emissions from United States Natural Gas Gathering and Processing. Environ. Sci. Technol. 49, 10718–10727. https://doi.org/10.1021/acs.est.5b02275

Michanowicz, D.R., Buonocore, J.J., Rowland, S.T., Konschnik, K.E., Goho, S.A., Bernstein, A.S., 2017. A national assessment of underground natural gas storage: identifying wells with designs likely vulnerable to a single-point-of-failure. Environ. Res. Lett. 12, 064004. https://doi.org/10.1088/1748-9326/aa7030

Mitchell, A.L., Tkacik, D.S., Roscioli, J.R., Herndon, S.C., Yacovitch, T.I., Martinez, D.M., Vaughn, T.L., Williams, L.L., Sullivan, M.R., Floerchinger, C., Omara, M., Subramanian, R., Zimmerle, D., Marchese, A.J., Robinson, A.L., 2015. Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results. Environ. Sci. Technol. 49, 3219–3227. https://doi.org/10.1021/es5052809

Moore, C.W., Zielinska, B., P?tron, G., Jackson, R.B., 2014. Air Impacts of Increased Natural Gas Acquisition, Processing, and Use: A Critical Review. Environ. Sci. Technol. 48, 8349–8359. https://doi.org/10.1021/es4053472

Nathan, B.J., Golston, L.M., O'Brien, A.S., Ross, K., Harrison, W.A., Tao, L., Lary, D.J., Johnson, D.R., Covington, A.N., Clark, N.N., others, 2015. Near-field characterization of methane emission variability from a compressor station using a model aircraft. Environ. Sci. Technol. 49, 7896–7903.

National Academies of Sciences, Medicine, 2018. Improving characterization of anthropogenic methane emissions in the United States. National Academies Press.

Neumann, P.P., Asadi, S., Lilienthal, A.J., Bartholmai, M., Schiller, M., 2011. Micro-drone for wind vector estimation and gas distribution mapping. J. IEEE Robot. Autom. Mag. 6, 1–11.

Nisbet, E.G., Manning, M.R., Dlugokencky, E.J., Fisher, R.E., Lowry, D., Michel, S.E., Myhre, C.L., Platt, S.M., Allen, G., Bousquet, P., 2019. Very strong atmospheric methane growth in the 4 years 2014–2017: Implications for the Paris Agreement. Glob. Biogeochem. Cycles 33, 318–342.

Omara, M., Sullivan, M.R., Li, X., Subramanian, R., Robinson, A.L., Presto, A.A., 2016. Methane emissions from conventional and unconventional natural gas production sites in the Marcellus Shale Basin. Environ. Sci. Technol. 50, 2099–2107.

Omara, M., Zimmerman, N., Sullivan, M.R., Li, X., Ellis, A., Cesa, R., Subramanian, R., Presto, A.A., Robinson, A.L., 2018. Methane emissions from natural gas production sites in the United States: Data synthesis and national estimate. Environ. Sci. Technol. 52, 12915–12925.

Palacios, V., George, S.R., von Fischer, J.C., Mohlin, K., 2017. Integrating Leak Quantification into Natural Gas Utility Operations. Environmental Defense Fund.

Patel, P., 2017. Monitoring Methane. ACS Publ. 3, 679–682.

Peckham, S.D., Hutton, E.W., Norris, B., 2013. A component-based approach to integrated modeling in the geosciences: The design of CSDMS. Comput. Geosci. 53, 3–12.

Peischl, J., Eilerman, S.J., Neuman, J.A., Aikin, K.C., de Gouw, J., Gilman, J.B., Herndon, S.C., Nadkarni, R., Trainer, M., Warneke, C., 2018. Quantifying methane and ethane emissions to the atmosphere from central and western US oil and natural gas production regions. J. Geophys. Res. Atmospheres 123, 7725–7740.

Phillips, N.G., Ackley, R., Crosson, E.R., Down, A., Hutyra, L.R., Brondfield, M., Karr, J.D., Zhao, K., Jackson, R.B., 2013. Mapping urban pipeline leaks: Methane leaks across Boston. Environ. Pollut. 173, 1–4. https://doi.org/10.1016/j.envpol.2012.11.003

Providence Photonics, 2018. Quantitative optical gas imaging with GL320. URL https://www.providencephotonics.com/leak-quantification (accessed 10.21.18).

Ravikumar, A.P., Brandt, A.R., 2017. Designing better methane mitigation policies: the challenge of distributed small sources in the natural gas sector. Environ. Res. Lett. 12, 044023. https://doi.org/10.1088/1748-9326/aa6791

Ravikumar, A.P., Roda-Stuart, D., Liu, R., Bradley, A., Bergerson, J., Nie, Y., Zhang, S., Bi, X., Brandt, A.R., 2020. Repeated leak detection and repair surveys reduce methane emissions over scale of years. Environ. Res. Lett. 15, 034029.

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

Ravikumar, A.P., Wang, J., Brandt, A.R., 2017. Are Optical Gas Imaging Technologies Effective For Methane Leak Detection? Environ. Sci. Technol. 51, 718–724. https://doi.org/10.1021/acs.est.6b03906

Ravikumar, A.P., Wang, J., McGuire, M., Bell, C.S., Zimmerle, D., Brandt, A.R., 2018. Good versus Good Enough? Empirical tests of methane leak detection sensitivity of a commercial infrared camera. Environ. Sci. Technol.

Rebellion Photonics, 2018. Gas cloud imaging camera. URL https://rebellionphotonics.com/gci-camera.html (accessed 10.21.18).

Reese, D., Melvin, C., Sadik, W., 2006. Smart LDAR: Pipe Dream or Potential Reality?

Rella, C.W., Tsai, T.R., Botkin, C.G., Crosson, E.R., Steele, D., 2015. Measuring Emissions from Oil and Natural Gas Well Pads Using the Mobile Flux Plane Technique. Environ. Sci. Technol. 49, 4742–4748. https://doi.org/10.1021/acs.est.5b00099

Rhodes, B.C., 2011. PyEphem: astronomical ephemeris for Python. Astrophys. Source Code Libr.

Riris, H., Numata, K., Wu, S., Gonzalez, B., Rodriguez, M., Kawa, S., Mao, J., 2017. Methane measurements from space: technical challenges and solutions, in: Turner, M.D., Kamerman, G.W. (Eds.), Laser Radar Technology and Applications XXII. Presented at the SPIE Defense + Security, p. 1019104. https://doi.org/10.1117/12.2262516

Ro, K.S., Johnson, M.H., Hunt, P.G., Flesch, T.K., 2011. Measuring trace gas emission from multi-distributed sources using vertical radial plume mapping (VRPM) and backward Lagrangian stochastic (bLS) techniques. Atmosphere 2, 553–566.

Robertson, A.M., Edie, R., Snare, D., Soltis, J., Field, R.A., Burkhart, M.D., Bell, C.S., Zimmerle, D., Murphy, S.M., 2017. Variation in Methane Emission Rates from Well Pads in Four Oil and Gas Basins with Contrasting Production Volumes and Compositions. Environ. Sci. Technol. 51, 8832–8840.

Rojas, A.J., Gonzalez, L.F., Motta, N., Villa, T.F., 2015. Design and flight testing of an integrated solar powered UAV and WSN for remote gas sensing. IEEE, pp. 1–10. https://doi.org/10.1109/AERO.2015.7119209

Roscioli, J.R., Yacovitch, T.I., Floerchinger, C., Mitchell, A.L., Tkacik, D.S., Subramanian, R., Martinez, D.M., Vaughn, T.L., Williams, L., Zimmerle, D., others, 2015. Measurements of methane emissions from natural gas gathering facilities and processing plants: measurement methods. Atmospheric Meas. Tech. 8, 2017.

Roy, A.A., Adams, P.J., Robinson, A.L., 2014. Air pollutant emissions from the development, production, and processing of Marcellus Shale natural gas. J. Air Waste Manag. Assoc. 64, 19–37. https://doi.org/10.1080/10962247.2013.826151

Saunier, S., Haugland, T., Pederstad, A., 2014. Quantifying cost-effectiveness of systematic leak detection and repair programs using infrared cameras (No. CL-13–27). Carbon Limits.

Schneising, O., Burrows, J.P., Dickerson, R.R., Buchwitz, M., Reuter, M., Bovensmann, H., 2014. Remote sensing of fugitive methane emissions from oil and gas production in North American tight geologic formations: Remote sensing of

fugitive methane emissions from oil and gas production. Earths Future 2, 548–558. https://doi.org/10.1002/2014EF000265
Schwietzke, S., Harrison, M., Lauderdale, T., Branson, K., Conley, S., George, F.C., Jordan, D., Jersey, G.R., Zhang, C., Mairs, H.L., 2018. Aerially-guided leak detection and repair: A pilot field study for evaluating the potential of methane emission detection and cost-effectiveness. J. Air Waste Manag. Assoc.
Schwietzke, S., Pétron, G., Conley, S.A., Pickering, C., Mielke-Maday, I., Dlugokencky, E.J., Tans, P.P., Vaughn, T., Bell, C.S., Zimmerle, D., others, 2017. Improved mechanistic understanding of natural gas methane emissions from spatially-resolved aircraft measurements. Environ. Sci. Technol.
Smith, B., John, G., Christensen, L.E., Chen, Y., 2017. Fugitive methane leak detection using sUAS and miniature laser spectrometer payload: System, application and groundtruthing tests, in: Unmanned Aircraft Systems (ICUAS), 2017 International Conference On. IEEE, pp. 369–374.
Stocker, T., 2014. Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press.
Subramanian, R., Williams, L.L., Vaughn, T.L., Zimmerle, D., Roscioli, J.R., Herndon, S.C., Yacovitch, T.I., Floerchinger, C., Tkacik, D.S., Mitchell, A.L., Sullivan, M.R., Dallmann, T.R., Robinson, A.L., 2015. Methane Emissions from Natural Gas Compressor Stations in the Transmission and Storage Sector: Measurements and Comparisons with the EPA Greenhouse Gas Reporting Program Protocol. Environ. Sci. Technol. 49, 3252–3261. https://doi.org/10.1021/es5060258
Szulczynski, B., Gebicki, J., 2017. Currently Commercially Available Chemical Sensors Employed for Detection of Volatile Organic Compounds in Outdoor and Indoor Air. Environments 4, 21. https://doi.org/10.3390/environments4010021
Tadić, J.M., Michalak, A.M., Iraci, L.T., Ilic, V., Biraud, S.C., Feldman, D.R., Bui, T.V., Johnson, M.S., Loewenstein, M., Jeong, S., others, 2017. Elliptic cylinder airborne sampling and geostatistical mass balance approach for quantifying local greenhouse gas emissions. Environ. Sci. Technol.
Tao, L., Sun, K., Miller, D.J., Pan, D., Golston, L.M., Zondlo, M.A., 2015. Low-power, open-path mobile sensing platform for high-resolution measurements of greenhouse gases and air pollutants. Appl. Phys. B 119, 153–164.
Terry, C., Argyle, M., Meyer, S., Sander, L., Hirst, B., 2017. Mapping methane sources and their emission rates using an aircraft. Lead. Edge 36, 33–35. https://doi.org/10.1190/tle36010033.1
Thoma, E., Squier, B., 2014. OTM 33 Geospatial Measurement of Air Pollution, Remote Emissions Quantification (GMAP-REQ) and OTM33A Geospatial Measurement of Air Pollution-Remote Emissions Quantification-Direct Assessment (GMAP-REQ-DA). US Environmental Protection Agency, Cincinnati, OH.
Thorpe, A.K., Frankenberg, C., Thompson, D.R., Duren, R.M., Aubrey, A.D., Bue, B.D., Green, R.O., Gerilowski, K., Krings, T., Borchardt, J., 2017. Airborne DOAS retrievals of methane, carbon dioxide, and water vapor concentrations at high spatial resolution: application to AVIRIS-NG. Atmospheric Meas. Tech. 10, 3833.
Townsend-Small, A., Tyler, S.C., Pataki, D.E., Xu, X., Christensen, L.E., 2012. Isotopic measurements of atmospheric methane in Los Angeles, California, USA: Influence

of fugitive fossil fuel emissions. J. Geophys. Res. Atmospheres 117. https://doi.org/10.1029/2011JD016826
Varon, D.J., Jacob, D.J., McKeever, J., Jervis, D., Durak, B.O., Xia, Y., Huang, Y., 2018. Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. Atmospheric Meas. Tech. 11, 5673–5686.
Vaughn, T.L., Bell, C.S., Pickering, C.K., Schwietzke, S., Heath, G.A., Pétron, G., Zimmerle, D.J., Schnell, R.C., Nummedal, D., 2018. Temporal variability largely explains top-down/bottom-up difference in methane emission estimates from a natural gas production region. Proc. Natl. Acad. Sci. 201805687.
Verhulst, K.R., Karion, A., Kim, J., Salameh, P.K., Keeling, R.F., Newman, S., Miller, J., Sloop, C., Pongetti, T., Rao, P., others, 2017. Carbon dioxide and methane measurements from the Los Angeles Megacity Carbon Project: 1. Calibration, urban enhancements, and uncertainty estimates. Atmos Chem Phys Discuss.
von Fischer, J.C., Brewer, P.E., Chamberlain, S., Gaylord, A., von Fischer, J., 2016. Mobile natural gas leak surveys indicate that two utilities have high false negative rates, in: AGU Fall Meeting Abstracts.
von Fischer, J.C., Cooley, D., Chamberlain, S., Gaylord, A., Griebenow, C.J., Hamburg, S.P., Salo, J., Schumacher, R., Theobald, D., Ham, J., 2017. Rapid, Vehicle-Based Identification of Location and Magnitude of Urban Natural Gas Pipeline Leaks. Environ. Sci. Technol. 51, 4091–4099. https://doi.org/10.1021/acs.est.6b06095
Wainner, R.T., Aubut, N.F., Laderer, M.C., Frish, M.B., 2017. Scanning, standoff TDLAS leak imaging and quantification, in: SPIE Commercial+ Scientific Sensing and Imaging. International Society for Optics and Photonics, pp. 1021006–1021006.
Weller, Z., Roscioli, J.R., Daube, W.C., Lamb, B.K., Ferrara, T., Brewer, P.E., von Fischer, J.C., 2018. Vehicle-based methane surveys for finding natural gas leaks and estimating their size: validation and uncertainty. Environ. Sci. Technol.
Yacovitch, T.I., Daube, C., Vaughn, T.L., Bell, C.S., Roscioli, J.R., Knighton, W.B., Nelson, D.D., Zimmerle, D., Pétron, G., Herndon, S.C., 2017. Natural gas facility methane emissions: measurements by tracer flux ratio in two US natural gas producing basins. Elem Sci Anth 5.
Yacovitch, T.I., Herndon, S.C., Pétron, G., Kofler, J., Lyon, D., Zahniser, M.S., Kolb, C.E., 2015. Mobile laboratory observations of methane emissions in the Barnett Shale region. Environ. Sci. Technol. 49, 7889–7895.
Yang, S., Talbot, R., Frish, M., Golston, L., Aubut, N., Zondlo, M., Gretencord, C., McSpiritt, J., 2018. Natural Gas Fugitive Leak Detection Using an Unmanned Aerial Vehicle: Measurement System Description and Mass Balance Approach. Atmosphere 9, 383.
Yen, C.-H., Horng, J.-J., 2009. Volatile organic compounds (VOCs) emission characteristics and control strategies for a petrochemical industrial area in middle Taiwan. J. Environ. Sci. Health Part A 44, 1424–1429.
Zavala-Araiza, D., Alvarez, R.A., Lyon, D.R., Allen, D.T., Marchese, A.J., Zimmerle, D.J., Hamburg, S.P., 2017. Super-emitters in natural gas infrastructure are caused by abnormal process conditions. Nat. Commun. 8, 14012. https://doi.org/10.1038/ncomms14012

Zavala-Araiza, D., Herndon, S.C., Roscioli, J.R., Yacovitch, T.I., Johnson, M.R., Tyner, D.R., Omara, M., Knighton, B., 2018. Methane emissions from oil and gas production sites in Alberta, Canada. Elem Sci Anth 6.

Zavala-Araiza, D., Lyon, D.R., Alvarez, R.A., Davis, K.J., Harriss, R., Herndon, S.C., Karion, A., Kort, E.A., Lamb, B.K., Lan, X., others, 2015. Reconciling divergent estimates of oil and gas methane emissions. Proc. Natl. Acad. Sci. 112, 15597–15602.

Zazzeri, G., Lowry, D., Fisher, R.E., France, J.L., Lanoisell?, M., Nisbet, E.G., 2015. Plume mapping and isotopic characterisation of anthropogenic methane sources. Atmos. Environ. 110, 151–162. https://doi.org/10.1016/j.atmosenv.2015.03.029