Techno-economic assessment of options to mitigate super-emitters in the natural gas supply chain

Primary Supervisor: Dr. Adam Hawkes

Co-Supervisors: Dr. Paul Balcombe and Dr. Jasmin Cooper

Supervised student: Ratanon Suksumrun CID: 01541619

Department of Chemical Engineering

Abstract

Super-emitters are infamously known for emitting high amounts of methane in the oil/gas industry. They are stochastic by nature as methane emissions vary in the natural gas supply chain by numerous factors, hence it is difficult to predict their existence.

In this study, real-world methane emissions are collected to evaluate which methane emissions are classified as super-emitters. Just collecting emissions data is not enough, data needs to be reviewed on how, when and where it was collected, type of emissions and its underlying source. Methane emissions that followed the aforementioned criteria is ideal for efficient analysis of the super-emitters.

A modelling workflow for this study is outlined to construct a high-resolution probabilistic emissions model. From data collection, data wrangling, algorithm implementation, data visualisation and analysis. Initial steps of the model are used to characterise methane emissions with statistical distributions to appropriately quantify the distribution's heavy-tail, in which the heavy-tail characteristic indicates a super-emitter.

A Monte Carlo simulation was then performed to produce 10,000 curves of methane emissions as a percentage of the EUR (Estimated Ultimate Recovery). In which the EUR is also fitted with a distribution of the same type as the methane emissions' best fitted distribution. The incentive for the Monte Carlo simulation is to statistically signify that a low EUR and a high methane emission influences the existence of the super-emitters significantly.

In-depth analysis of the super-emitters is conducted by constructing Lorenz curves to identify the skewness within the data. Results showed that majority of the datasets showed significant skewness, with average emissions by the top 5% emitters were on the lower end of the emission range. However, the average emissions did not skew the data as immensely as expected, but the presence of a few top 5% emitters with magnitudes larger than the average emissions were the main influence in skewness contribution.

Economic assessment was conducted to evaluate low-cost, methane monitoring/reduction technologies against a reduced emissions scenario. Results from the assessment show that mitigation costs were negative, indicating the amount of money being saved rather than spent. An optimisation model is implemented and results showed that there is virtually no cost required to mitigate a tonne of CO_2 eq. Sensitivity analysis is performed on the natural gas price and how it affects the mitigation costs in 10 years. The analysis indicated that the natural gas price is likely to increase relative to the current gas price and this increases the cost-effectiveness of the innovative mitigation technologies. This demonstrates that low-cost methane monitoring/reduction technologies are efficient and cost-effective.

Acknowledgements

I would like to express my gratitude to Dr Adam Hawkes for giving me the opportunity to undertake this MSc project. I would also like to express my sincere gratitude to Dr Paul Balcombe and Dr Jasmin Cooper for their supervision, guidance and continual support throughout the whole project. Not only that, both of you have given me full control of the project and I managed to gain new skills and knowledge in this process.

I would also like to thank the companies and research institutes that took their time to provide me some data input for this study. Thank you to Yori Jamin from Clearstone Engineering, Joshua Anhalt and Mike D'Antoni from GreenPath Energy for providing me their methane emissions datasets. Thank you to Hendrik Hamann and Levente Klein from IBM and James Scherer from Aeris Technologies for providing me economic information for the economic assessment of this study. Thank you to Evar Umeozor from CERI for providing me the ICERM database and guidance on the report "Economic and Environmental Impacts of Methane Emissions Reduction In The Natural Gas Supply Chain". Without all this information, this study would not be possible.

Lastly, I would like to express my sincere thanks to my girlfriend Cat, my family and all my friends, who have conversed with me about the project, for their continual and moral support.

Contents

List of Tables

List of Figures

Nomenclature

Symbols/Abbreviations	Description	Unit
AIC	Akaike's information criterion	-
AIMS	An Intelligent Methane Management Solution	-
ARPA-E	Advanced Research Projects Agency–Energy	-
BU	Bottom-Up	-
BIC	Bayesian Information Criterion	-
CAD	Canadian Dollars	[$CAD]
CAPEX	Capital Expenditure	[$USD] [$CAD]
CCAC	Climate and Clean Air Coalition	-
CERI	Canadian Energy Research Institute	-
CEPEI	Canadian Energy Partnership for Environmental Innovation	-
CH_4	Methane	[m^3/day] [g] [m^3] [gCO_2 eq / MJ-HHV] [kg/h] [% of EUR]
CO_2	Carbon Dioxide	[gCO_2 eq / MJ-HHV]
CO_2 eq	Carbon Dioxide Equivalent	-
E	Rate of absolute methane emissions	[kg/h]
EDF	Environmental Defense Fund	-

EPA	Environmental Protection Agency	-
EUR	Estimated Ultimate Recovery	[Mm^3]
f_e	Proportional loss rate	-
GHG	Greenhouse Gas	-
GHGI	Greenhouse Gas Inventory	-
GJ	Gigajoules	-
GWP	Global Warming Potential	[gCH_4/ gCO_2] [$tonneCH_4$/ $tonneCO_2$]
HHV	Higher Heating Value	-
i	Emission source *i*	-
IBM	International Business Machines	-
ICF	Inner City Fund	-
IR	Infra-Red	-
j	Mitigation technology *j*	-
LB	Lower bound	-
LDAR	Leak Detection and Repair	-
LDC	Local Distribution Companies	-
LNG	Liquefied Natural Gas	-
MONITOR	Methane Observation Networks with Innovative Technology to Obtain Reductions	-
MC_j	Mitigation cost for technology *j*	[$CAD/tonne CO_2 eq]
OPEX	Operating Expenditures	[$USD] [$CAD]
P_r	Rate of methane production	[kg/h]
REC	Reduced Emissions Completions	-

RE_i	Reduced methane emission for emission source *i*	[m^3/day]
$RE_{Max,i}$	Maximum reduced methane emission for emission source *i*	[m^3/day]
SE AB	Southeast Alberta	-
TC	TransCanada	-
TD	Top-Down	-
TDLAS	Tunable Diode Laser Absorption Spectroscopy	-
THC	Total Hydrocarbon	[THC m^3/h]
UB	Upper bound	-
USD	US Dollars	[$USD]
VRU	Vapour Recovery Units	-
α	Sites with f_e below the 85th percentile	-
β	Super-emitting sites where f_e exceeds the 85th percentile	-
γ	Super-emitting sites where f_e exceeds maximum rate of absolute emissions	-

Chapter 1: Introduction

Methane (CH_4) is a prominent greenhouse gas and its emissions are frequently reported in the natural gas supply chain of the upstream field in the oil/gas industry, alongside CO_2 emissions. Methane contribution to the total emissions was approximated to be 58% of CO_2's radiative forcing effects (0.97W/m^2 for methane and 1.68 W/m^2 for CO_2). However, methane's immediate emission is 120 times larger than CO_2, meaning that a gram of methane contributes 120 times more than a gram of CO_2 to warm the Earth's atmosphere (Balcombe *et al.*, 2018). Reducing methane emissions not only prevents global warming from occurring but also prevent an increase in mortality rates, various GHG emissions (eg.CO_2), water contamination as well as resources depletion (CCAC, 2016) (Mandler and Allison, 2018).

However, the challenge to reduce methane emissions is strenuous because natural gas is made up of 78-93% methane (Skone, Littlefield and Marriott, 2011), hence, methane emissions are quite prevalent. In addition, the range of methane emissions vary broadly due to differences in estimation methods and this spurred controversy across literature (Alvarez *et al.*, 2018). Estimation methods are prone to uncertainties due to lack of data, assumptions, methodology differences and variations (such as different technologies, well types, regional policies, etc.) in the natural gas supply chain (Balcombe *et al.*, 2015). With such uncertainties and flexible approaches, there is no absolute standard to tackle this problem.

To increase the complexity of the problem, a phenomenon called super-emitters exist in the supply chain. Super-emitters are described as a group of emitters, whether it is an equipment or facility-specific emitter, that emit an extremely high amount of methane. Thus, contributing to the increase in total emissions in the oil/gas sector (Balcombe *et al.*, 2015). Additionally, super-emitters can change from one equipment/facility to another, this makes it more difficult to predict its existence. Currently, it is desirable to identify solutions for mitigating super-emitters due to its immense contribution to the total emissions.

The chosen geographical location for this study is Alberta, Canada. This location was chosen since the oil/gas industry in Alberta is known to have the largest source of methane emissions out of all the industry sectors. In 2014, 31.4 MT of CO_2 eq were emitted from the oil/gas industry alone. This value alone is 70% of the provincial CH_4 emissions and 25% of all GHG emissions from the upstream sector (Government of Alberta, 2019).

Recent news from the EDF showed that large amount of methane emissions in Canada were severely underreported by the oil/gas industries. With 60 oil/gas production sites in Red Deer were shown to emit methane 15 times larger than usual reports by industry. Nelson and Zavala-Araiza indicated that the contribution to such high methane emissions were the super-emitters (Nelson, 2018). It is evident that with Alberta's natural gas infrastructure as well as the underestimated methane emissions reported by industry, super-emitters are a prevalent problem that must be tackled immediately. To understand a gain deeper understanding of these super-emitters, the following procedures are performed:

A general literature review (not specific to Alberta, Canada) was carried out to set the context of the super-emitters. This includes discussing the emission sources within the natural gas supply chain, characterisation of the methane emissions, the definition of super-emitters, the current mitigating technologies for methane emissions and the research project's objectives. Following on the literature review, the methodology of this research project was outlined: a technical assessment was conducted to identify super-emitters along with as reducing the emissions in a cost-effective way via economic assessment. After the assessments, a section dedicated to the results are discussed further for insights on the super-emitters. Finally, a conclusion in this report is to summarise the existence of super-emitters, identifying them and the possible solutions of reducing them.

Chapter 2: Literature Review

2.1 Emission sources within the natural gas supply chain

This section provides information on the components of the natural gas supply chain as well as discussing the main emission sources to give an overview of the methane emissions. This includes all the attainable routes of the natural gas production from site preparation on different well types, processing natural gas and transmission of the gas to consumers. The possibility of the inclusion of the LNG (Liquified Natural Gas) route is considered (Liquefaction, Storage and Transport and Regasification blocks). The supply chain is illustrated in Figure 1.

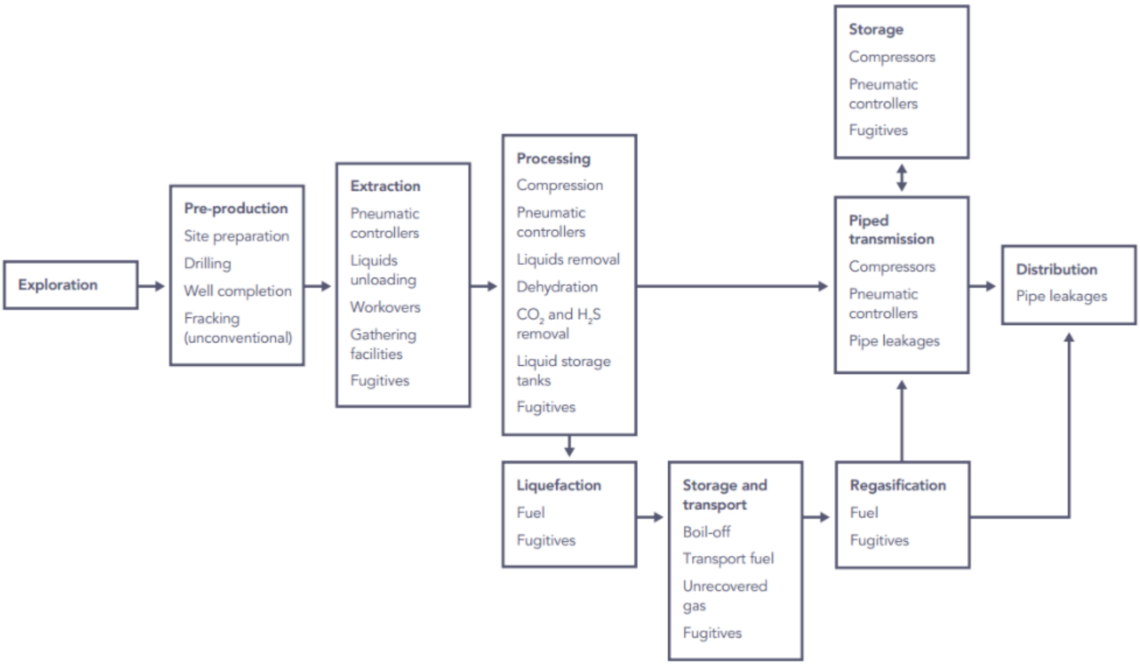

Figure 1: Diagram of the full natural gas supply chain with individual components for characterisation of methane emissions. Source: (Balcombe *et al.*, 2017)

The natural gas supply chain starts by an oil platform exploring a feasible gas well based on geological data. Once a gas well has been found and dependent on the well type, the platform drills into the well and prepares the well site for production. The extraction (production) phase removes the natural gas from the well and sent to a processing plant for purification. If the natural gas needs to be transferred in high quantities, it is sent to the LNG route. After processing and regasification, the natural gas is sent to a piped transmission and distribute to the consumers or storage.

The total estimated supply chain emissions of methane and carbon dioxide were found to be in the range of 2 and 42 gCO_2 eq / MJ-HHV, with some higher estimated emissions (45 - 99 gCO_2 eq / MJ-HHV) from specific stages shown in Figure 33 and 34 in Appendix section. Note that total estimated supply chain emissions are defined as individual total estimates from 250 literatures. The two highest emission sources are from well completions and liquids unloading processes as shown by the frequency of the circles in Figure 33 and 34. A review on these two sources will be extensively discussed in this section due to its frequent emissions. The graphs indicated that most of the total greenhouse gas emissions are located at the lowest part of the range, supported by each of the stages' median estimates. (Balcombe *et al.*, 2017).

Wells completion is a series of operations that involves insertion of a casing into the drilled hole followed by cement injection to wrap around the casing in preparation of natural gas production. Fracking was employed for unconventional wells during wells completion and significant methane emissions were from the flowback of the fracturing fluid, estimated at 0-87 gCO_2 eq / MJ-HHV. An emissions data for well completions is shown in Figure 35 in Appendix section.

From the data in Figure 35, unconventional wells that utilised RECs (Reduced Emissions Completions) equipment are substantially lower than wells with non-RECs equipment. RECs are equipment that captures the flowing natural gas during flowback process for purification and transmit to the product pipeline or vent the gas if it is deemed unusable. Conventional well types emissions are relatively small as there was no need for fracking as well as RECs equipment (Balcombe *et al.*, 2017). Figure 36 in Appendix section shows a schematic of a REC equipment.

It is evident that using REC type equipment substantially reduced the methane emissions. Utilising RECs would signify that wells completion is not a possible super-emitter source, however, it is important to take this into account when identifying super-emitters in unconventional wells scenarios and that it is an ideal mitigating solution in the supply chain (Balcombe, Brandon and Hawkes, 2018).

Liquids unloading is considered a very high methane emissions source based on Figure 33 and 34. The purpose of this stage is to remove liquids that accumulated at the bottom of the well bore due to either a decrease in pressure, the ratio of gas to liquid or the gas velocity as a cause of well maturity (David T Allen *et al.*, 2015). The associated process/equipment for liquids unloading: well blowdown, well swabbing, plunger lift systems, foaming agents, velocity tubing and an artificial lift system. The range of methane emissions for liquids unloading is extremely large (at about 0 – 500000 m^3CH_4/year) due to the variation of gas wells requirements. Some wells do require unloading while others do not, depending on the characteristics of the gas wells.

Most frequent emissions data for liquids unloading are for plunger lift systems and well blowdowns with a calculated total 273,568MT ($9.30x10^{12}$ gCO_2 eq / MJ-HHV)of emitted methane from 60,810 wells based on (OAQPS, 2014)'s Table 2-4. The data was conducted on US regions that are categorised in NEMS (National Energy Modelling System), so this will be used to produce EIA's Annual Energy Outlook. The table indicated that the emissions for plunger lifts and well blowdowns vary substantially due to inconsistent well samples for each unloading process in different regions. Hence, there are uncertainties in data comparison, but it is evident that the amount of methane emitted from the equipment are significant.

To further support the above statement, (David T Allen *et al.*, 2015) conducted a more detailed investigation 107 gas wells in US regions on liquids unloading particularly in the use of automatic/manual plunger lift systems and without the plunger lift systems. A graph on the estimated methane emissions from these gas wells are shown in Figure 2 and some results are tabulated in Table 1.

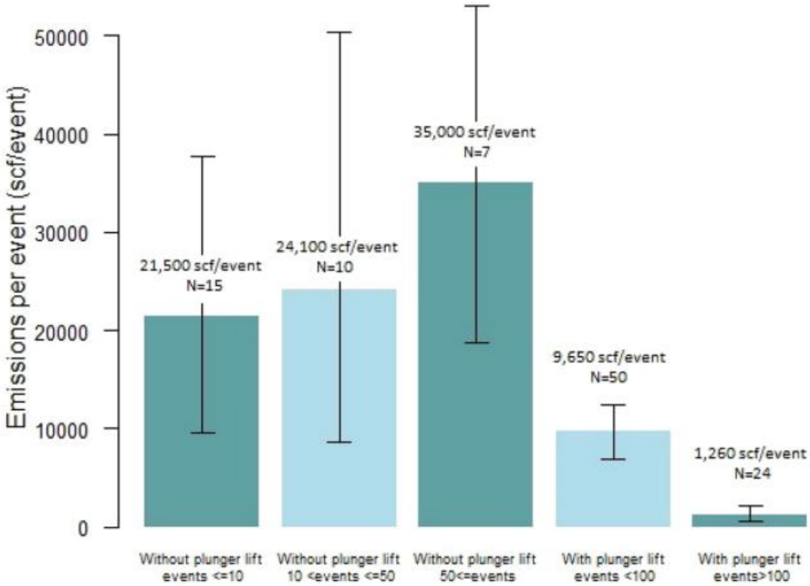

Figure 2: Methane emissions in standard cubic feet per liquids unloading event on 107 gas wells per year. Source: (David T Allen et al., 2015)

Table 1: Tabulated results of methane emissions from liquids unloading per event per year in m^3 and g of CH4 and time range in minutes. Adapted from (David T Allen et al., 2015)

Unloading Type	**Sampled Wells**	**Emissions (m^3 of CH_4)**	**Emissions gCH_4**	**Time range (mins)**	**Time range (hours)**
Without plunger lift systems	32	595 – 991	11000 – 2600000	10.2 – 270	0.17-4.5
Manual plunger lift systems	50	6 - 1388	4000-940000	1.8 - 180	0.03 – 3
Automatic plunger lift systems	25	1.42 - 227	1000 - 150000	1 – 20	0.02 – 0.33

Figure 2 stated that the mean emissions for the 32 sampled wells without plunger lift systems are 609m^3, 683m^3, 991m^3 per event. This applies to 15 wells that require less than 10 unloading, 10 wells that require 11-50 unloading times and 7 gas wells that require more than 51 unloading times respectively. 50 gas wells with plunger lifts are categorised as manual systems had mean emission at 273m^3 per event. 25 gas wells with automated plunger lifts were sampled with more than one event unlike the previous wells, the number of sampled events were in the range of 2 – 70 events. This produced methane emissions of 1.42 – 227m^3 (1000 – 150000g methane) per event. Due to the variable frequency of the events, a mean emission of 36m^3 (1260 scf) per event was used for suitable comparison in Figure 2.

It would be concluded that automatic plunger lift systems emit the highest amount of methane, however, Balcombe et al stated that it is invalid to conclude this statement (Balcombe *et al.*, 2017). This is because the wells that do not require unloading and require manual unloading are sampled as one event whilst automatic plunger systems are sampled at more than one event. There are numerous factors that could affect such a high frequency of event, this could be due to the characteristics/age of the well or how the automated plunger lift system was designed. Further assessment on these factors is ideal due to the high variability of the emissions and the fact that it is one of the significant key methane emissions sources as well as a possible super emitter source. Quantifying this as a national estimate, annual emissions from wells with plunger lifts system alone are at 190Gg CH_4 with 80% of these emissions are associated with venting more than 100 times per year. On average per year, emission from wells that require venting more than 100 times are at 27Mg CH_4 with 95% confidence intervals at 10 – 50 Mg CH_4 (Allen *et al.*, 2013).

2.2 Characterisation of methane emissions

General characterisation methods are categorised as bottom-up (BU) and top-down (TD) methods. BU methods involve estimating emissions at a regional or national level, estimations were done by directly measuring the source of emissions and aggregate them to the desired level. TD methods aggregate emissions by utilising airborne equipment such as satellites, tower networks or aircrafts. This method allowed measurements to be made across multiple regions and landscapes more conveniently (Alvarez *et al.*, 2018).

To identify super-emitters, it would be best to utilise the BU method as it tackles directly into the methane emissions source. TD methods lack the clarity on identify super-emitters as it does not directly investigate the underlying source of emissions. However, both methods are associated with their own uncertainties, so it is difficult to do a comparison. BU methods can lead to underestimated values due to exclusion of emission sources or overlooked super-emitters. (Alvarez *et al.*, 2018) compared the BU method results with the EPA's 2015 GHGI report and concluded that the GHGI had overlooked the super-emitters factor, resulting in an underestimated emission in the GHGI. TD methods can lead to overestimated values due to the complexity of assigning methane concentrations in the atmosphere to certain emission sources (Balcombe *et al.*, 2017).

One of the widely used BU methods was adopting statistical methods to create probabilistic distribution emission models followed by a Monte Carlo simulation. Utilising the Monte Carlo simulation results in random generation of draws of the probability distributions of the total supply chain emissions. It can then be assessed which scenario contributes to higher or lower emissions. A study by (Balcombe, Brandon and Hawkes, 2018) adopted this BU method and produced a cumulative distribution graph of the total supply chain of methane emissions in Figure 3.

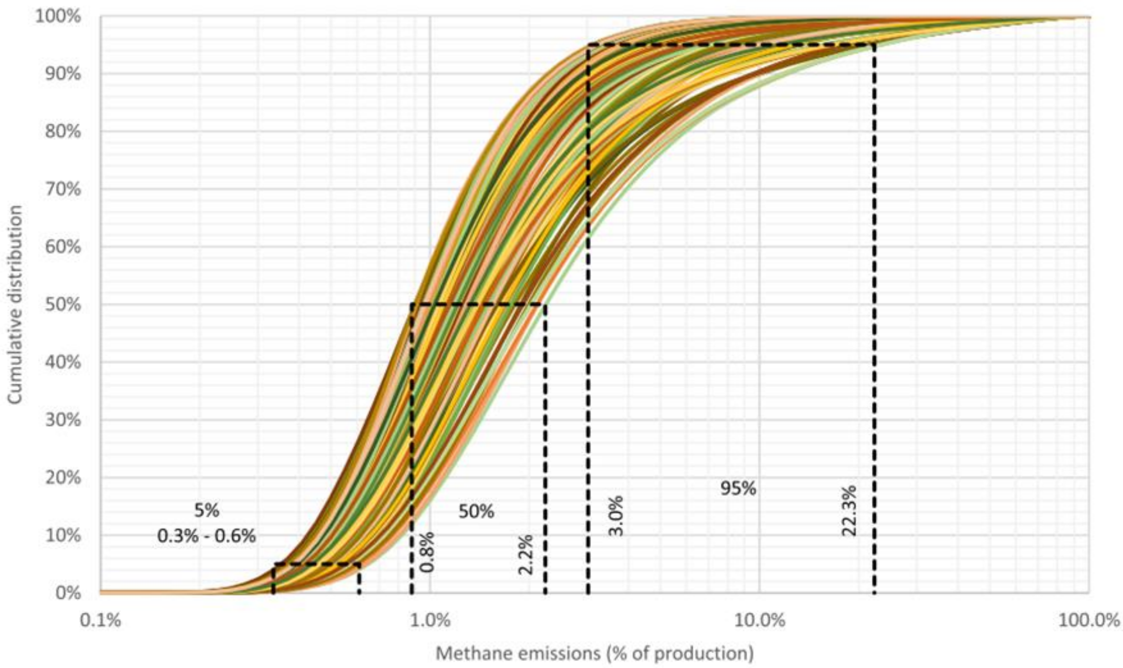

Figure 3: Cumulative distribution graph of the methane emissions as the total supply chain emissions. Source: (Balcombe, Brandon and Hawkes, 2018)

Note that the emissions data in this study utilised global emissions data and not specific in regions.160 supply chain scenarios were simulated, and total emissions were estimated 10,000 times for each scenario to form cumulative distribution curves. Methane emissions are expressed as percentages of an estimated ultimate recovery of a well (EUR), which is essentially the total production of a well. Key results from the graph stated that 5th and 95th percentile was ranged at 0.3-0.6% and 3.0 – 22.30% respectively. Converting this into CO_2 equivalents using the GWP100, emissions are at 1.6-3.2 gCO_2eq/gCH_4 (5th percentile) and 16.2 – 120.2 gCO_2eq/gCH_4 (95th percentile) (Balcombe, Brandon and Hawkes, 2018).

(Balcombe, Brandon and Hawkes, 2018) discussed about identifying the skewness of the distributions, which includes ratio of the mean and median emissions for each supply chain scenario and the log-logistic distributions. The study stated that the average ratio in the 160 scenarios were 2.3 and typically in the range of 1.2-4.6. A high ratio indicated that super-emitters are key contributors to the total supply chain emissions. The probability distribution curve in Figure 3 is best described as a log-log-logistic distribution, which is used to significantly characterise the “heavy-tail” of a distribution. These “heavy-tails” are often an indication of the existence of super-emitters.

In addition, the Monte Carlo simulation also explored the scenario with the implemented mitigating technologies, these include utilising RECs equipment, manual plunger lifts systems and pipeline distributions made of plastic (Balcombe, Brandon and Hawkes, 2018). This method is ideal for the research project as it identifies the existence of super-emitters as well as simulating various supply chain scenarios that incorporate diverse mitigating technologies.

(Zavala-Araiza *et al.*, 2017) took a different approach by identifying a threshold for super-emitters of 26kg CH_4 /hr per production site. The approach Zavala-Araiza et al took was using component-based emissions along with Monte Carlo simulations to see how these emissions contribute to a production site to be classified as a site-level super emitter. The component-based emissions are then compared with the site-based emissions study by (Zavala-Araiza, D. R. Lyon, *et al.*, 2015). The site-based emissions were characterised by statistical estimators (Zavala-Araiza, D. R. Lyon, *et al.*, 2015). A classification of super-emitter in the study was developed in Figure 4.

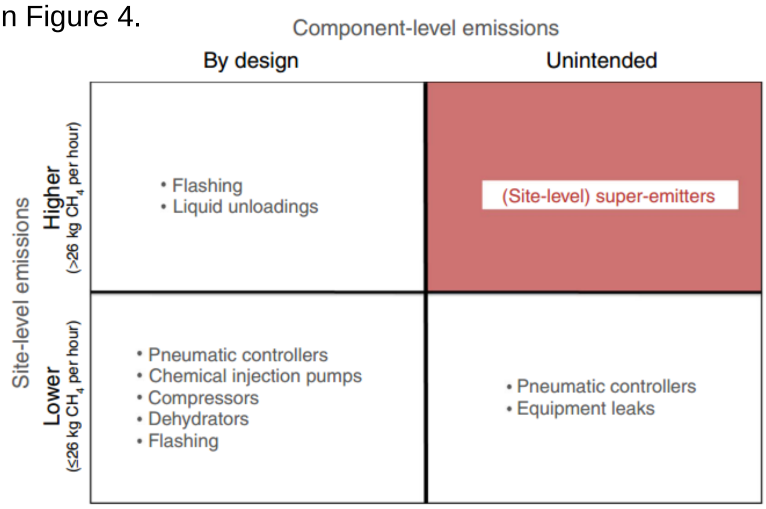

Figure 4: Grouping of site-level emissions and component-level emissions. Source: (Zavala-Araiza *et al.*, 2017)

The results from the study (Zavala-Araiza *et al.*, 2017) stated that 13 out of 170 sites in the Barnett Shale region exhibited site-level super-emitters behavior. The Monte Carlo simulation results displayed that cumulative component-based emissions (20,000 kgCH_4 /hr) were 1.5 times lower than the site-based emissions (31,000 kgCH_4 /hr). The highest emitting sites from the site-based distribution showed cumulative emissions 9 times more than the highest emitting sites using component-based methods. The stark difference between the site-based and component-based methods determined the existence of the super-emitters and concluded that they arise from abnormal process conditions. The two studies by Zavala-Araiza et al highlighted the methodological differences as well as identifying a threshold supports the identification of super-emitters. Using a threshold would be ideal in defining super-emitters in the research project, however, comparing two BU methods to identify the difference in values as super-emitters would prove to be complex as well as introduce more uncertainties.

Another study by Zavala-Araiza et al proposed an alternative way to define super-emitters as a function of proportional loss rates (Zavala-Araiza, D. Lyon, *et al.*, 2015). Super-emitters were characterised to have a proportional loss rate above a specific threshold. Thresholds can be assigned dependent on the author of the methane emissions study. Equation 1 is used to model the proportional loss rate of methane emissions.

$$E = f_e P_r \ (1)$$

Where E is the rate of absolute methane emissions (kg/h), $f_e$is the proportional loss rate and P_r is the rate of methane production (kg/h).

For (Zavala-Araiza, D. Lyon, *et al.*, 2015)'s case, the threshold was at the 85th percentile within the range of gas production groups in the study. The study categorised production sites to α,β and γ. α are classified as sites with f_e below the 85th percentile, β and γ are classified as super-emitting sites where f_e exceeds the 85th percentile and f_e reached the maximum rate of absolute emissions respectively. The purpose of this method was to introduce an alternative quantitative assessment of methane emissions for production sites subsequent to the development of statistical data. This identification method of super-emitters would be ideal to implement in the project due to its systematic approach, however, this method focused on varying natural gas production rates instead of varying the supply chain routes. This method would be out of the scope of this project but still a considerable method to implement.

2.3 Super-emitters and mitigation technologies

Super-emitters are abnormalities that arises when certain equipment or process conditions emit an extremely high amount of methane. In literature, super-emitters are discovered via a highly skewed distribution profile of emissions, resulting in a "heavy-tail" distribution. This indicated there are more excessive methane emissions sources than expected (Brandt *et al.*, 2014).

Multiple studies have frequently reported the existence of super-emitters from fugitive and venting emissions (Brandt, Heath and Cooley, 2016) (Zavala-Araiza, D. Lyon, *et al.*, 2015). Fugitive (unintentional) emissions such as leaks, equipment malfunctions, inefficient operating procedures and maintenance policies are frequently found in sources such as the extraction phase, gathering stations, processing plants, transmission, compressors and distribution stages (EPA, 2006a). While venting emissions occurs from plunger systems in liquids unloading, venting in compressors and pneumatic devices (David T. Allen *et al.*, 2015). Liquids unloading itself is also considered to be a super-emitter (Balcombe *et al.*, 2015).

This phenomenon is responsible for the wide range and large magnitude of CH_4 emissions. The behaviour of the super-emitters are classified as stochastic and dynamic (Zavala-Araiza *et al.*, 2017). This meant that super-emitters can occur at any stage within the supply chain and/or at any moment in time. There is no absolute procedure to remove the super-emitters, but it is possible to mitigate their occurrence. More attention has been given on identifying cost-effective mitigation solutions as super-emitters are becoming more prevalent and currently is an active area of research.

ICF has compiled specific technologies to tackle the super-emitters sources, it is noteworthy that there are more mitigation technologies provided by the EPA, but ICF specifically focuses on the technologies for the largest methane sources in the natural gas supply chain. Technologies included are: plunger lift systems in liquids unloading, centrifugal compressors with dry seals or with wet seal degassing recovery systems, vapour recovery units on storage tanks, low-bleed pneumatic devices, rod packing replacement in reciprocating compressors, electric pumps and LDAR (Leak Detection and Repair) programs (ICF International, 2014). A cost benefit-analysis for the technologies are summarised in Table 15 in Appendix section and a reduction cost curve in Figure 5.

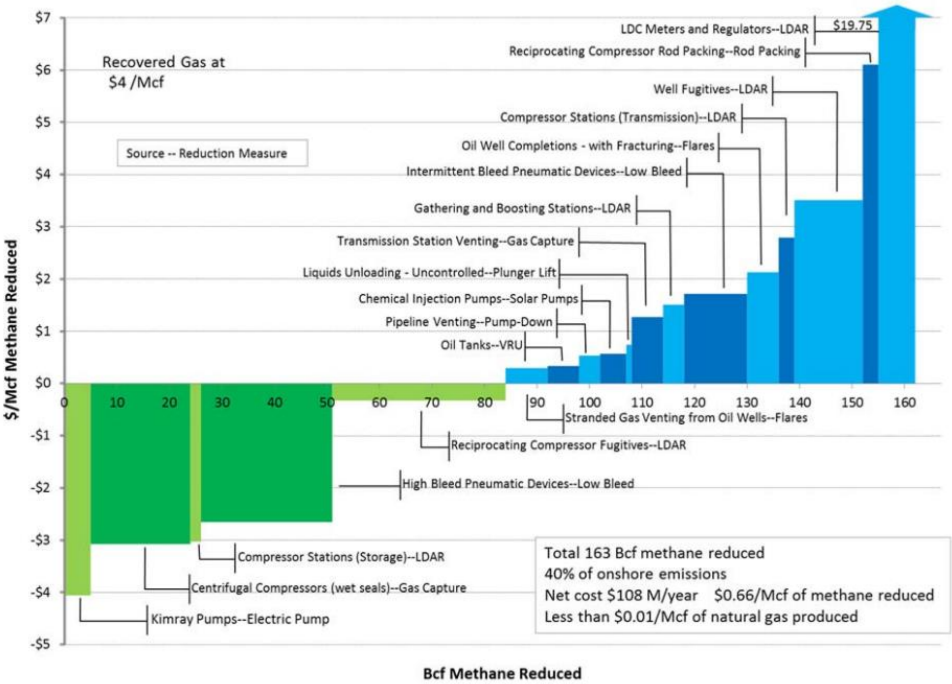

Figure 5: Reduction cost curve in 2018 scenario relative to the EPA 2011 Inventory data. Dollars are quantified as in millions Source: (ICF International, 2014)

Note the US Dollars in this section were from the Year 2013 (ICF International, 2014). Figure 5 stated a total reduction of 4.6 billion m^3 CH_4, equivalent to 40% of methane onshore emissions with the current technologies across the natural gas supply chain. Net total cost was $0.66/Mcf of methane reduced and annualised cost was at $108M/year. Total annual cost was the annual savings of \$164 million and annual cost of \$272 million, based on the total areas of the green and blue bar respectively. The assumptions associated with this marginal cost curve are the credit of a recovered natural gas is \$4/Mcf along with a 10% capital cost and discount rate for calculations of \$/Mcf methane reduced for each technology. The capital costs for the implementation of these technologies are estimated to be at \$2.2 billion.

ICF recommended to start implementing the LDAR (Leak Detection and Repair) program as the fugitive emissions across the supply chain represents the 36% of the emission reduction potential. Within these 36%, the sources are fugitive emissions from compressor stations, wells, gathering stations and LDC facilities. Along with replacement from high-bleed to low-bleed pneumatic devices, reduction in venting and installing recovery systems in wet seal compressors. Each of these represent 30%, 22% and 12% of the emission reduction potential respectively (ICF International, 2014).

LDAR is a program that utilises leak detection equipment such as IR (infra-red) camera to identify the leaks and repair them using various procedures. EPA improved the LDAR program by adopting the U.S. EPA Method 21, where leakage detections are done in a quarterly, monthly or yearly basis. Repairs are advised when leakages exceed a certain range (generally at 10000ppm) (EPA, 2006a). LDAR is the most recognised solution to mitigate methane emissions from leaks so this would be considered as one of the solutions to use in this research project. The CAPEX (based on Table 15) for the LDAR is at \$124,000 and its percentage reduction potential is at 60% (assuming quarterly checks), however, the OPEX is significantly one of the cheapest and has 100% methane abatement efficiency relative to other technologies. If the LDAR assumes monthly basis checks, the reduction potential would increase to 80% (ICF International, 2014).

Pneumatic-type controllers is used to maintain temperature, pressure and liquid level within the processes in the supply chain. The operations are dependent on the pressure of gas flow as well as discharging a small proportion of gas. The amount of discharged gas is dependent on the type of pneumatic devices, categorised as low-bleed ($0.17m^3/hr$) and high bleed ($0.85m^3/hr$). In addition, there are intermittent bleed types and discharge rate is between 0.17 and 0.85 m^3/hr but only discharge gas during actuation.

It is desirable to convert the high-bleed controllers to low-bleed to reduce emissions. However, high-bleed controllers are compulsory for high-speed valves in compressors, so this is not possible for chains requiring compressors. The conversion to low-bleed controllers would be applicable to 60% of high-bleeds in the LNG, storage and transmission chain. Along with 80% within the processing chain and 90% in the rest of the chain requiring pneumatic controllers. The CAPEX for the conversion ranges from \$260 - \$2300 dependent on the time of conversion or for retrofit purposes. Payback period is within the range between 4 to 27 months and with a reduction potential of over 90%. This is a considerable option to mitigate super-emitters, however, the methane abatement efficiency is up to 84% and subject to change dependent on the discharge rate of the high-bleed controllers. Factors affecting discharge rate includes the gas pressure, equipment age and the amount of actuations (ICF International, 2014) (EPA, 2006c).

Venting reduction technologies include kimray pumps conversion to electric pumps and installing vapour recovery unit systems. Kimray pumps are used to separate water from the natural gas, the pumps are operated by natural gas' mechanical energy and then vent the used natural gas at the end of operation. Electric pumps reduce the venting by operating with electricity, so its methane abatement efficiency is at 100%. However, implementing the electric pumps would adopt solar panels, generators or grid connections to the pump can propose a challenge. In addition, should solar panels be installed, the unpredictability of the weather will impose another obstacle (ICF International, 2014) (EPA, 2006d). A feasible and considered option with relatively cheap CAPEX and OPEX but little revenue compared to other technologies.

Vapour recovery unit (VRU) systems reduce venting by capturing 95% flashing vapours during natural gas storage as these vapours would be vented, increasing methane emissions. Methane abatement efficiency is up to 95% but its operations can be affected with the compressor and scrubber in the system. However, this technology is the most expensive in terms of CAPEX and OPEX, as well as the longest payback time. Nevertheless, VRUs have the highest revenue since the recovered vapours are sold (or further use as fuels for existing facilities) (ICF International, 2014) (EPA, 2006b).

Degassing recovery systems for wet seals compressors reduce methane emissions by recycling wet seals that captures escaping methane. The recycling options can be used as fuel for boilers and turbines, back to the compressor or vent it. The third option should be discarded to minimise any possible venting emissions. Methane abatement efficiency is up to 99% with relatively expensive CAPEX as well as a rewarding revenue and payback time of 3 months. With flexible recycling options to reduce compressor emissions, this would be another ideal mitigating solution (ICF International, 2014) (EPA, 2014).

From Section 2, it was discussed that utilising plunger lift systems is a possible super-emitter source whilst ICF suggested plunger lift systems would reduce methane emissions in liquids unloading. It is important to consider the study by (David T Allen *et al.*, 2015) that plunger lift systems can also contribute to the existence of super-emitters when conducting the techno-economic assessment.

Throughout the literature review, the natural gas supply chain was introduced and key emission sources such as wells completion and liquids unloading were discussed to understand the general overview of the supply chain emissions. The two sources were chosen due to its high frequent emissions across literature and how this would affect in the techno-economic aspects of the research project. Before the discussion on how to characterise methane emissions, a brief discussion on the use of GWP100 was tackled due to its controversy in quantifying methane emissions. In addition, alternative metric systems were introduced and concluded that how specific metric systems would impact the quality of the mitigation solutions. Two general methods, bottom-up and top-down, were used to identify methane emissions. BU methods were more widely used since it tackled emission sources more directly than TD methods. In addition, super-emitters identification methods were also extensively discussed. Finally, super-emitters were defined in this section and a detailed discussion on the mitigation technologies were also tackled. A few of the mitigation technologies were identified as they provided the best opportunity to reduce the largest methane emissions sources as stated by the economic analysis report by ICF (ICF International, 2014).

Super-emitters have been extensively researched in the USA (Allen *et al.*, 2013) (Zavala-Araiza *et al.*, 2017) (ICF International, 2014), however, there are less studies conducted in Canada. Consequently, methane emissions in Canada are underreported especially in Alberta, where the largest methane emissions are from the oil/gas industry (Nelson, 2018). EDF reported that recent methane emissions from the Red Deer region in Alberta based on a study by (Zavala-Araiza *et al.*, 2018). The study stated that the 12 of the 60 production sites in Red Deer are responsible for 75% of the methane emissions and concluded that super-emitters are the responsible for this. The study was comparable with North American's production sites due to similarities with Red Deer's production sites, however, it is incorrect to conclude that the emissions are representative of the Alberta region or Canada. The Government of Canada acted on this matter to implement methane regulations such as the LDAR program and tighter venting regulations for pneumatic devices, compressors, production facilities and wells completion for unconventional well types. The regulations will take place in the year of 2020 (2023 for venting regulations of production facilities and pneumatic devices) (Government of Canada, 2018). This is so that Canada can reach the 45% methane emission reduction by 2025. To achieve this, it is important to identify super-emitters and explore more mitigating options for these phenomena.

2.4 Project objectives and plan

The objective for this research project is to conduct a techno-economic assessment of the mitigation technologies to mitigate the existence of the super-emitters in the Alberta, Canada.

The first part of the project is to develop expertise in R, as this will be used to conduct statistical distribution models and initiate the Monte Carlo simulation later in the project. Whilst developing R skills, data on methane emissions in the Alberta region as well as various industrial practices and current mitigating technologies will then be collected. The collected data will then be used to develop the natural gas supply chain probabilistic emissions model. After the construction of the model, the Monte Carlo simulation will then be applied to assess the various supply chain scenarios. Each scenario can be used to identify which scenario contributes to the least and the most total methane emissions. The simulation will provide some robust information on the existence of super-emitters due to its stochastic nature. An economic assessment such as the cost-benefit analysis will be used to analyse the economics behind each mitigating technology. The technologies chosen are from ARPA-E's MONITOR program. This will then be used to develop recommendations on why the MONITOR technologies are better than conventional mitigating technologies along with the ideal operating practices. A preliminary schedule of the research plan is shown in Figure 6.

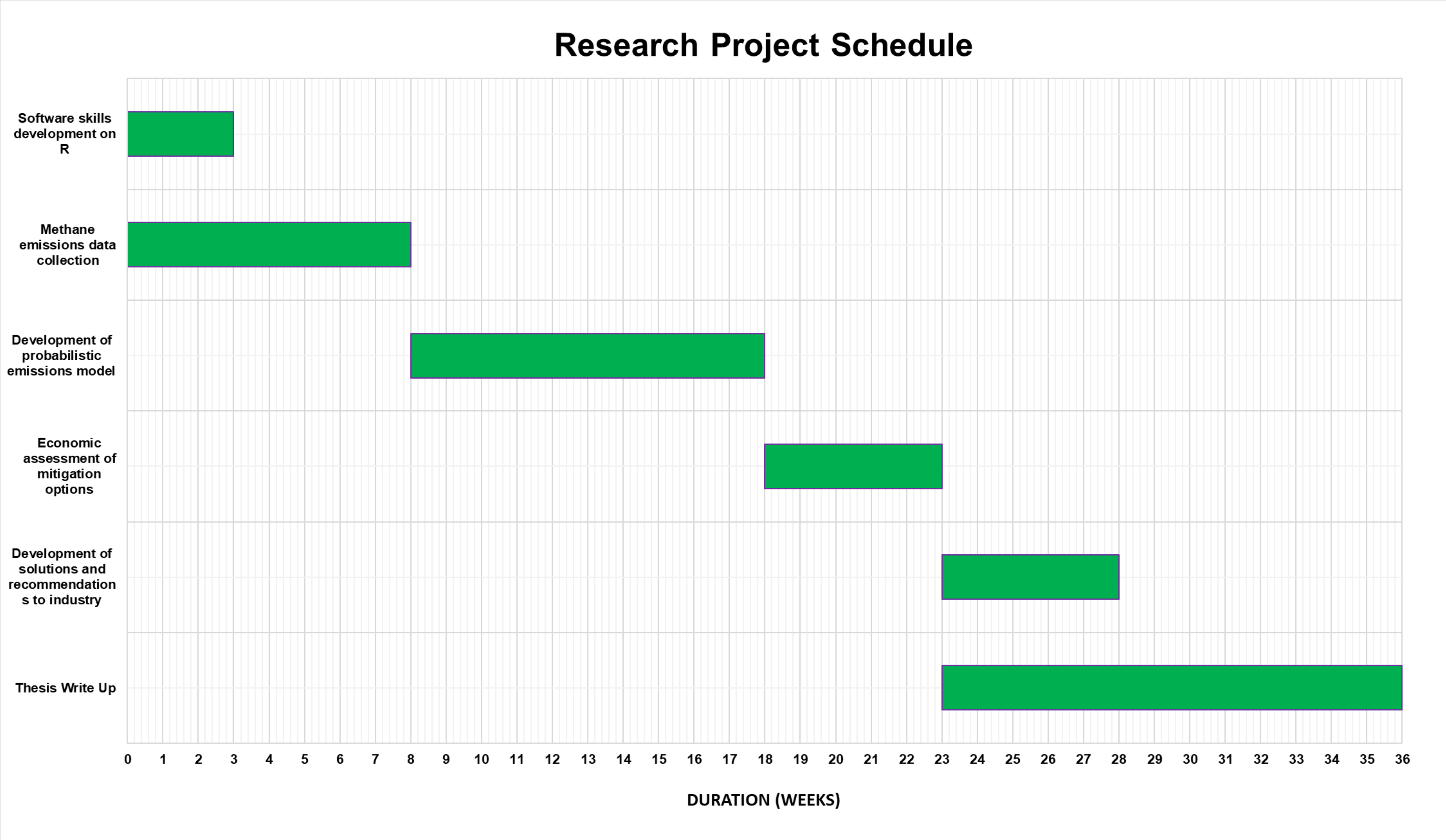

Figure 6: Gantt chart of the preliminary research project plan

Chapter 3: Techno-Economic Assessment Methodology

3.1 Data collection and pre-processing

The first part of the technical assessment of the project starts off by building a probabilistic emissions model for the collected Alberta's methane emission data. Due to the lack of accessible data in Canada, the considered stages are mainly from the upstream sector. The chosen stages in the upstream sector are gas wells, gas gathering networks, gas batteries, compressor stations. The natural gas supply chain for this study is shown below:

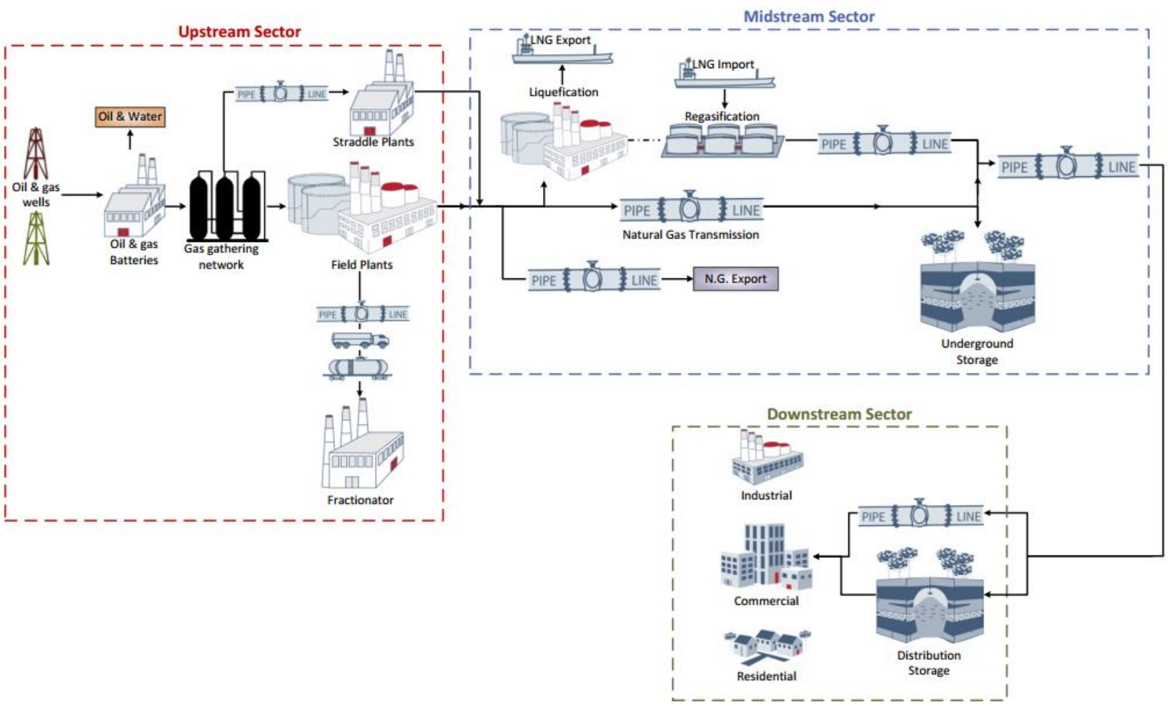

Figure 7: Natural Gas Supply Chain in Canada. Source: (Umeozor *et al.*, 2019)

The initial phase in identifying super-emitters efficiently is having a well-defined, robust data input for the probabilistic emissions model. Defining a robust data input for this study, the inputs should ideally display the emission source (equipment-specific such as compressors, separators, etc.) along with its emission category (the type of emission the source exhibits, venting or leaked emissions) and the source's reporting segment (the stages from the natural gas supply chain).

Two datasets from the upstream sector were provided by (Clearstone Engineering Ltd., 2018) and (GreenPath Energy Ltd., 2016). The focus of these datasets is that the emission type is exclusively leakage emission source whilst some venting emission sources were also included. These data were robust enough as it satisfied the previously mentioned criteria.

Two datasets from academia reports by (Zavala-Araiza *et al.*, 2018) and (O'Connell *et al.*, 2019) are also considered in this study; however, they are collected via airborne measurements. Due to the nature of its measurements, it will not be as compatible with the other considered datasets, as it does not underpin the root cause of the super-emitters. It only indicates that the super-emitter phenomenon does exist within the natural gas supply chain sites. The considered units of all methane emission sources in this study are in m^3/day. All data inputs are in the Appendix as Table 16, 17, 18 and 19.

Super-emitters are dynamic and stochastic by nature; hence the population and the emission value of the super-emitters interchange, dependent on numerous factors. One of the main factors that contribute to such high emission values is due to the estimated ultimate recovery (EUR) of a well, which is essentially the total lifetime methane production of a well. The EUR values vary based on different well reservoir types and these types include such as conventional gas, shale and coalbed methane. There is an incentive that by varying the values of the EUR, it contributes to the amount of methane being emitted and this was previously discussed in literature (Balcombe, Brandon and Hawkes, 2018). Further impacts made by the EUR to the methane emissions will be discussed in the results section. To investigate this further, methane emissions are modelled as a percentage of the EUR with the help of a Monte Carlo simulation via random number generation (Rubinstein and Kroese, 2011).

The purpose of conducting a Monte Carlo simulation is to predict the trend and statistically assess the existence of the super-emitters by randomly generate 10,000 estimates of methane emissions (% of EUR). 10,000 estimates were drawn from 10,000 randomly generated numbers from the collected EUR data whilst the methane emissions from the collected emission datasets are kept constant. 10,000 is the initial starting number of scenarios in this statistical assessment as the number of scenarios are large enough to counter the stochastic nature of the super-emitters. Hence it is the best methodology to identify super-emitters in the technical assessment (Balcombe, Brandon and Hawkes, 2019). However, due to the lack of accessible EUR data for Alberta, Canada, the considered EUR data in this study will be from the US as the proxy data and these were taken from (Balcombe *et al.*, 2015).The EUR data is shown in Table 20 in Appendix.

It is desirable that the proxy EUR data would be replaced with a more accurate Canadian EUR data. Nevertheless, the collected US EUR data displayed an accurate representation of the range of EUR values and compatible enough in this study. It is expected that the EUR will exhibit different values across different well reservoir types such as shale gas, coalbed methane and conventional natural gas in Canada. However, the impact of the EUR to the methane emissions with different well reservoir types is beyond the scope of this study, but nevertheless, it is desirable to conduct further research on this matter.

Once all data inputs were collected, an initial data wrangling phase was conducted for compatibility across datasets, suitable for RStudio to program algorithms, simulations and the probabilistic emissions model. After the data has been exported into RStudio, calculations and further data wrangling was also conducted, this includes manipulation of some categorical variables to comply with the natural gas supply chain stages for compatibility. An example for this was from the Clearstone dataset, where methane emissions' reporting segment are from "Gas Multiwell Effluent"," Gas Multiwell Group"," Gas Multiwell proration outside SE AB" and "Gas Single" facilities were changed to "Gas Battery" facilities, as all the mentioned facilities are a type of gas battery facilities. The considered units for methane emissions in this study are in m^3/day. From the Clearstone dataset, the emission rate obtained were in THC (Total Hydrocarbon) m^3/hour. This can be converted into methane emission rate by multiplying the original emission rate with the volume composition of methane in natural gas. The value for this is 92.87 vol% and complies with the CEPEI Methodology manual (Umeozor *et al.*, 2019). From the GreenPath dataset, emissions were provided as ft^3/min and was converted to m^3 by multiplying with 0.03515. A modelling workflow and a summary of the collected methane emissions are shown below in Figure 8 and Table 2 respectively.

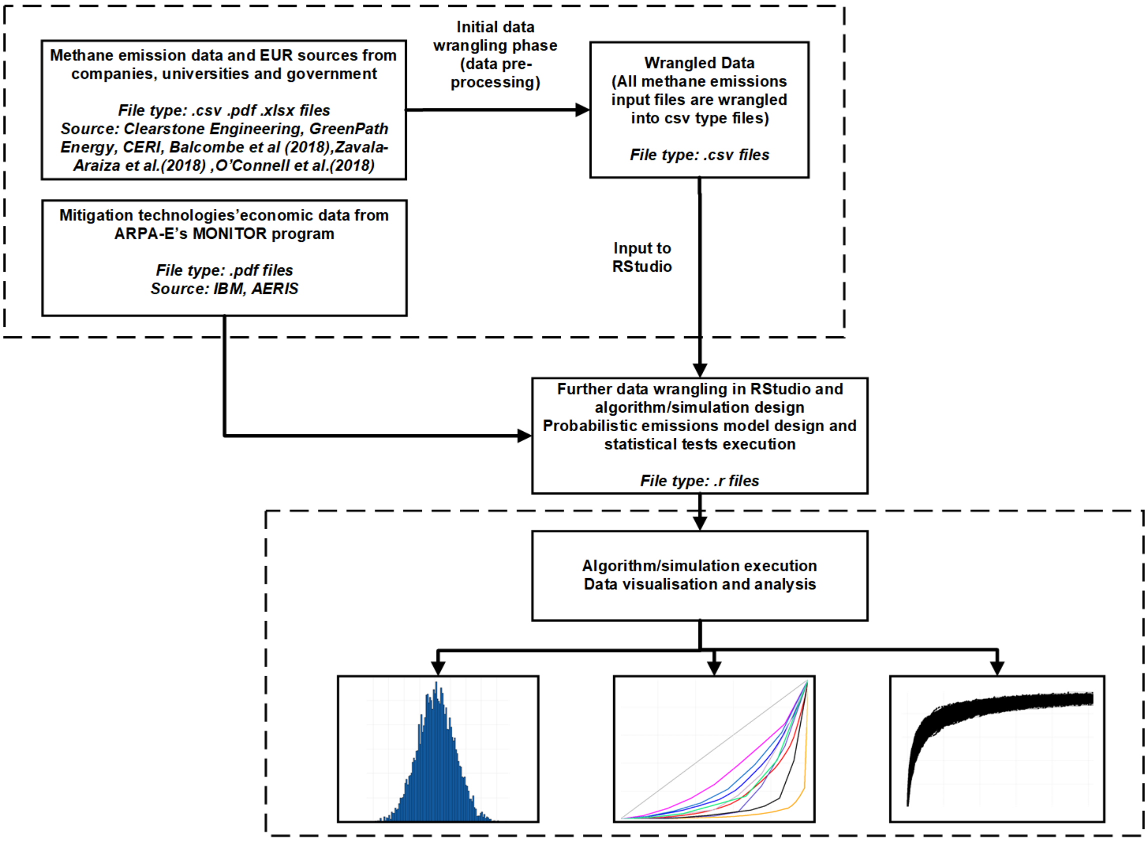

Figure 8: Methane emissions data modelling and pre-processing workflow

Table 2: List of emission sources along with its associated sector, the method, data points and reference sources

Reference	Sector/Method	Stage	Emission Source	Data Points
(Clearstone Engineering Ltd., 2018)	Upstream and Bottom-Up	Gas Batteries Compressor Stations Gas gathering systems	Reciprocating compressors Separators Wellheads Dehydrator-Glycols Flare knockout drums Production tanks (fixed-roof) Gas pipeline headers Screw compressors Electric-drive compressors Centrifugal compressors Pig traps Liquid pumps Line heaters	51 (Reciprocating compressors) 36 (Separators) 46 (Wellheads) 8 (Dehydrator-Glycols) 8 (Flare knockout drums) 13 (Production tanks) 7 (Gas pipeline headers) 8 (Screw compressors) 6 (Electric-drive compressors) 1 (Centrifugal compressors) 1 (Pig Trap) 1 (Liquid Pump) 2 (Line Heaters) 112 (Gas Batteries) 19 (Compressor stations) 64 (Gathering System)
(GreenPath Energy Ltd., 2016)	Upstream and Bottom-Up	Gas Batteries Gas Wells Compressor Station	Wellheads Filter/Separation (Separator) Compression (Compressors) Tankage	12 (Wellheads) 9 (Separator) 3 (Compression) 1(Tankage) 10 (Gas Batteries) 14 (Gas Wells) 1 (Compressor Station)
(Zavala-Araiza *et al.*, 2018)	Upstream and Top-Down	Gas Production	-	60
(O'Connell *et al.*, 2019)	Upstream and Top-Down	Gas Production	-	93
(Balcombe *et al.*, 2015)	-	EUR	-	53

For both Clearstone and GreenPath datasets, there are some emission sources and reporting segments that have a population range of 1-3. Due to the extreme small number of samples, they are not considered for the construction of the Lorenz curves in Section 3.4. However, some of these extremely small sources could be relevant in the analysis of the curves. Throughout this study, super-emitters will be defined as a small proportion of emitters (typically 20% or less) of the total population of emissions that emits at least 75% of the total methane emissions. However, the proportion of super-emitters can increase to more than 20% of total population. These proportions are then evaluated by exploring the magnitude of the top 5% emitters from its respective dataset. The top 5% emitters are defined to be the super-emitters, and this will be further discussed in Section 4.3.

3.2 Probabilistic emissions model

The goal of building a probabilistic emissions model is to model the best statistical distribution profile for each of the considered stages and its associated sources. The EUR data was also inputted into the model to find the best-fitted distribution. This step is crucial for the Monte

Carlo simulation as the EUR should ideally be from the same fitted distribution as the methane emissions for compatibility. The purpose of fitting a distribution is to investigate the nature of how the selected methane emissions are characterised and what inferences can be made from these distributions with regards to not just identifying super-emitters, but the nature of the methane emissions itself.

One of the advantages of RStudio are that it offers a broad range of packages specifically for statistical programming. One of the packages used in this study is called “fitdistrplus” package. The package offers various functions that fits data to a specific probability distribution. On top of that, the package also offers statistical tests such as the maximum log-likelihood and the goodness-of-fit tests (Delignette-Muller *et al.*, 2019). These quantitative tests, along with visualisation of the distribution, were used to statistically assess the best fitting distribution for each methane emission dataset. The considered distributions in this project are: log-normal, log-logistic, Weibull and Gamma. These distributions were considered as they are known to provide right skewed distributions as a basis for identifying super-emitters (Balcombe, Brandon and Hawkes, 2019) (Zavala-Araiza, D. R. Lyon, *et al.*, 2015) (Aitkin *et al.*, 2009).

In addition, the “actuar” package was also utilised as the functions related to the log-logistic distributions were not part of the “fitdistrplus” package. The “actuar” package offers heavy-tailed distribution functions and is extensively used within the actuarial science field (Goulet *et al.*, 2019). Nevertheless, the log-logistic distribution is appropriate enough to model methane emissions due to the distribution’s applicability for non-negative continuous random variables and produce moderately heavier-tails than the normal distribution (Aitkin *et al.*, 2009). The curves were estimated using the maximum log-likelihood estimation.

Using the aforementioned packages along with additional packages such as “dplyr” for convenient methane emissions data manipulation (Wickham, 2019) and “ggplot2” for elegant data visualisation (Wickham, 2016), tables detailing the distribution fitting analysis and graphs of the statistical distribution profile for each dataset are shown in Table 3,4,5 and Figure 9,10,11,12 and 13. Statistical tests will be further discussed in the results section.

Table 3: Maximum log-likelihood summary for the methane emission sources and EUR

	Maximum Log-Likelihood			
Source	Weibull	Gamma	Log-Normal	Log-Logistic
(Clearstone Engineering Ltd., 2018)	-546.4321	-587.3129	-518.0531	-511.2112
(GreenPath Energy Ltd., 2016)	-14.76842	-23.33093	-6.040871	-3.325216
(Zavala-Araiza *et al.*, 2018)	88.30029	68.99969	101.7549	104.4173
(O’Connell *et al.*, 2019)	-429.478	-434.5003	-415.755	-415.226
(Balcombe *et al.*, 2015)	-207.0892	-212.7005	-197.386	-194.8003

Table 4: Goodness-of-fit criteria for the methane emission sources and EUR

	Akaike's Information Criterion				Bayesian Information Criterion			
Source	Weibull	Gamma	Log-Normal	Log-Logistic	Weibull	Gamma	Log-Normal	Log-Logistic
(Clearstone Engineering Ltd., 2018)	1103.410	1178.626	1040.106	1026.422	1103.410	1185.172	1046.652	1032.968
(GreenPath Energy Ltd., 2016)	33.537	50.662	16.082	10.650	35.975	53.100	18.519	13.088
(Zavala-Araiza *et al.*, 2018)	-172.601	-133.999	-199.510	-204.835	-168.623	-130.021	-195.532	-200.857
(O'Connell *et al.*, 2019)	862.956	873.001	835.510	834.452	868.021	878.066	840.575	839.517
(Balcombe *et al.*, 2015)	418.178	429.401	398.772	393.601	421.289	432.512	401.883	396.771

Table 5: Goodness-of-fit statistical tests for the methane emission sources and EUR

	Cramer-von Mises Statistic				Anderson-Darling Statistic			
Source	Weibull	Gamma	Log-Normal	Log-Logistic	Weibull	Gamma	Log-Normal	Log-Logistic
(Clearstone Engineering Ltd., 2018)	1.091	3.286	0.234	0.181	7.146	17.017	1.929	1.564
(GreenPath Energy Ltd., 2016)	0.484	1.080	0.180	0.0897	2.842	5.317	1.159	0.640
(Zavala-Araiza *et al.*, 2018)	0.543	1.834	0.0965	0.0277	3.376	8.999	0.689	0.258
(O'Connell *et al.*, 2019)	0.364	0.624	0.0308	0.0168	2.486	3.516	0.237	0.148
(Balcombe *et al.*, 2015)	0.534	0.960	0.159	0.0632	3.268	5.046	1.205	0.615

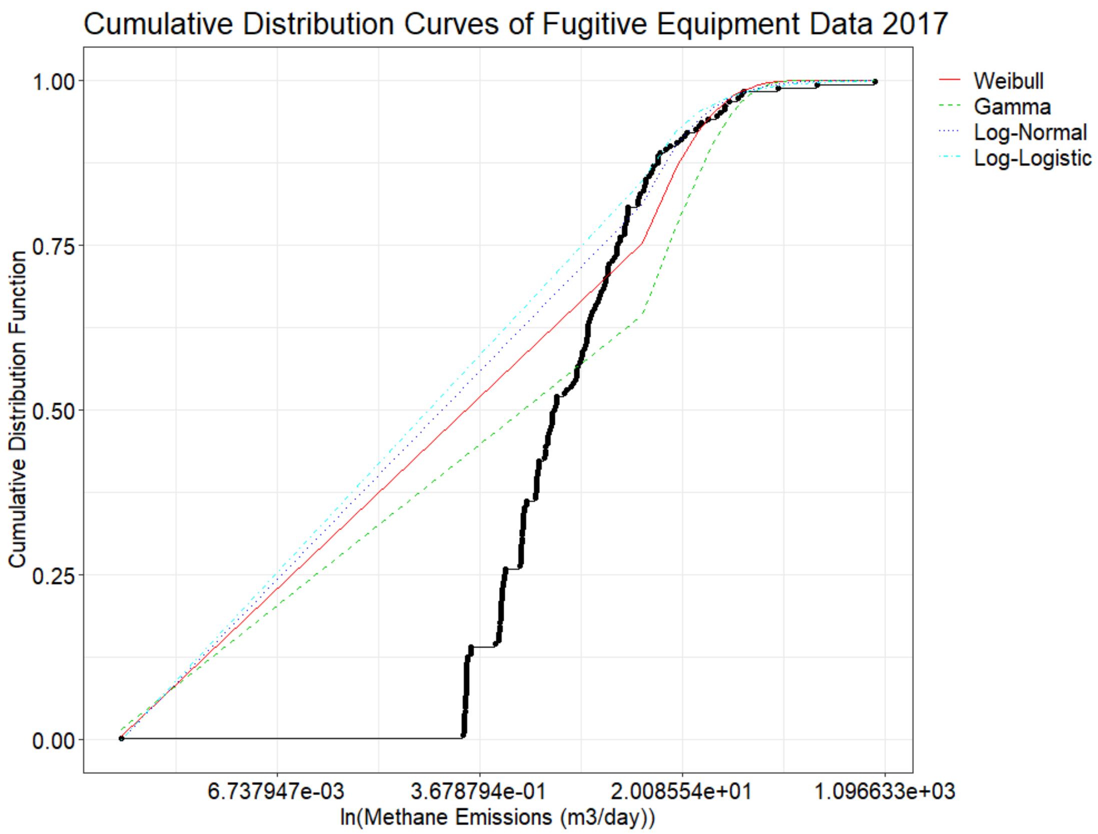

Figure 9: Weibull, gamma, log-normal, log-logistic cumulative distribution curve fitting for Clearstone

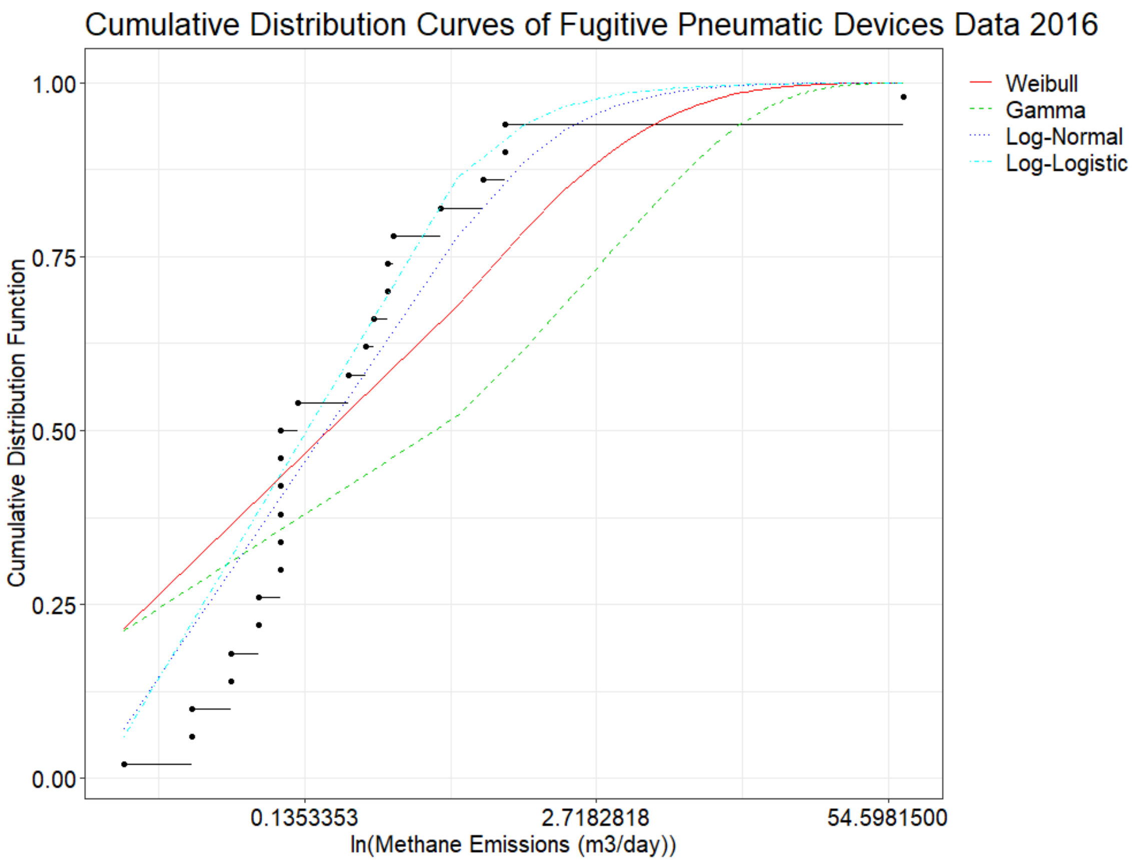

Figure 10: Weibull, gamma, log-normal, log-logistic cumulative distribution curve fitting for GreenPath

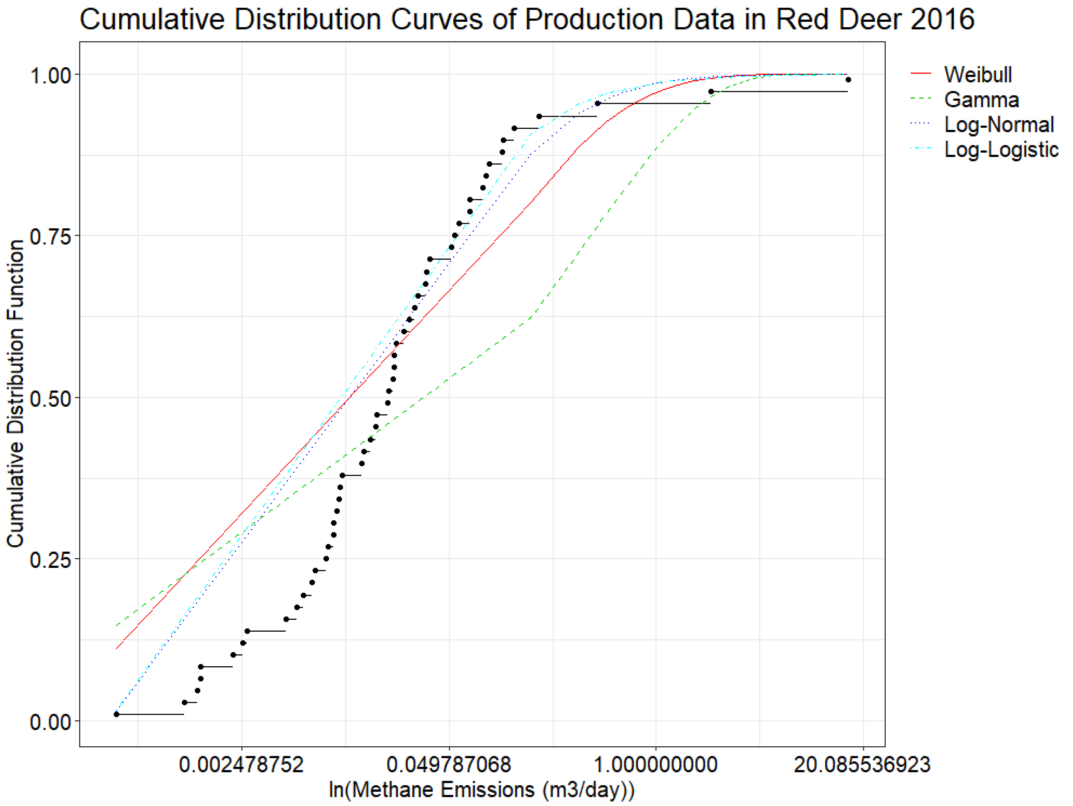

Figure 11: Weibull, gamma, log-normal, log-logistic cumulative distribution curve fitting for Zavala-Araiza

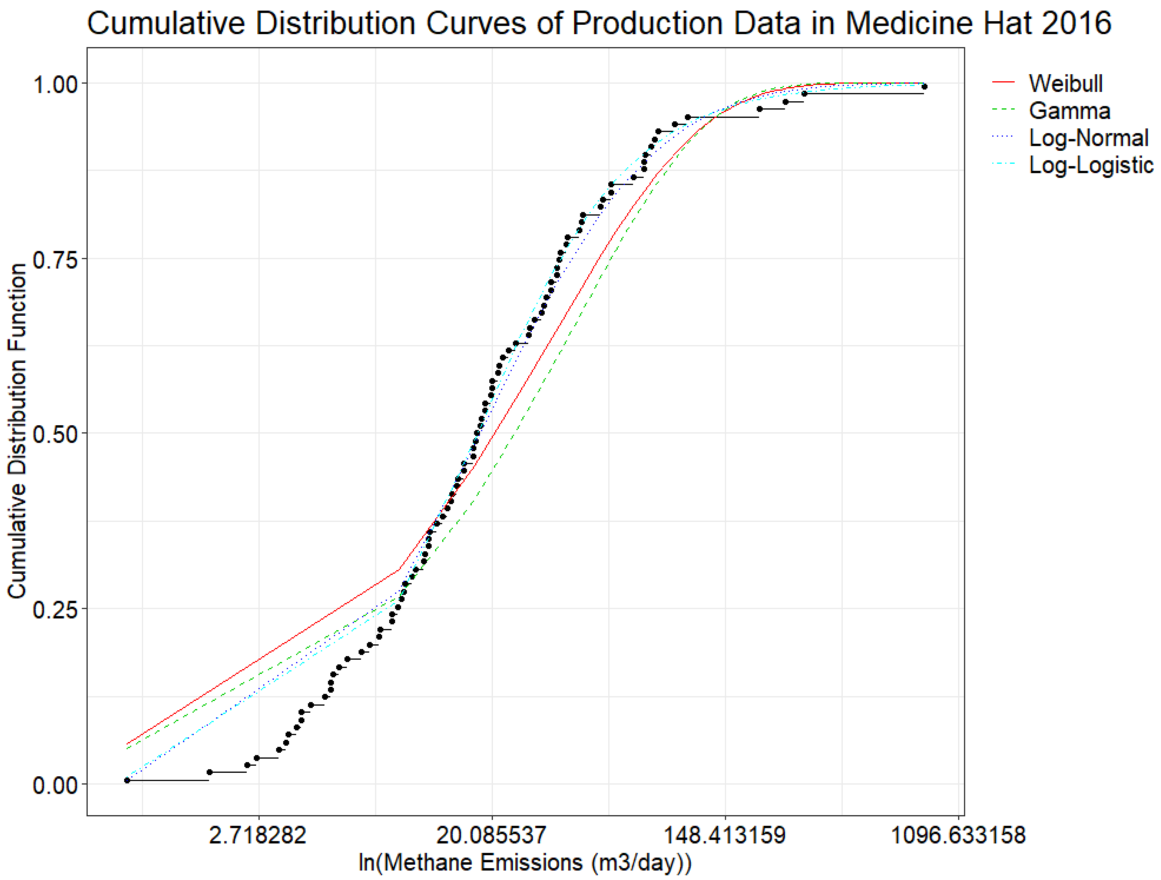

Figure 12: Weibull, gamma, log-normal, log-logistic cumulative distribution curve fitting for O'Connell

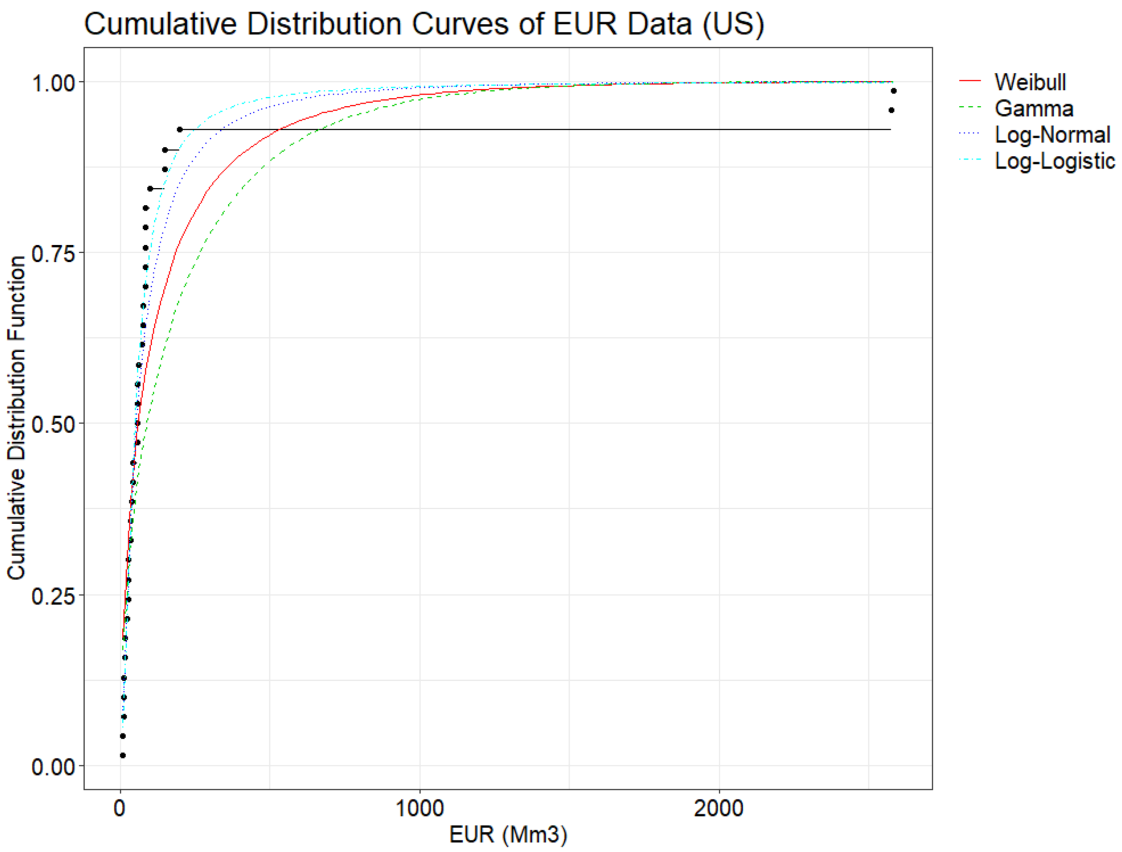

Figure 13: Weibull, gamma, log-normal, log-logistic cumulative distribution curve fitting for Balcombe EUR data

All cumulative distribution curves, except for the EUR data, were logged with a base of the exponential content. The purpose for this was to visualise the curves more clearly as the range of methane emissions were large and this almost caused the fitted distribution curves to coincide with each other. This would make it more difficult to compare the curves.

From the table of results and graphs, it can be concluded that for all methane emission sources and EUR are best fitted with a log-logistic distribution. There is a slight difference between the quantitative values of log-normal and log-logistic distribution from Table 3, 4 and 5. This indicated that the methane emission sources can also be described as log-normal. However, the nature of the logistic (and log-logistic) distributions are used in modelling logistic regressions and the required input values for the regression are categorical dependent variables. The methane emissions dataset used in this study are not categorical dependent variables as they do not depend on a certain category or discrete events for methane emissions to occur. In conclusion, it would be inappropriate to fit emissions with logistic distributions and should fit all emissions (including the EUR) with log-normal distributions.

3.3 Monte Carlo Simulation

After the construction of the probabilistic emissions model and evaluated the best fitted distributions, the model is then executed with a Monte Carlo simulation by generating 10,000 random numbers of the EUR from the log-normal distribution. The equation for modelling methane emissions as a percentage of the EUR is shown below:

$$Methane\ Emissions\ (\%\ of\ EUR) = \frac{Methane\ Emissions}{\ln(EUR) \sim N(4.005{,}1.259)}\ (2)$$

The emissions were expressed as a percentage of EUR are levelised emissions to consider any potential occurrence of the intermittent emissions in the supply chain such as liquids unloading or workovers. This methodology is a type of allocation method which is commonly used in the life cycle analysis to assess emitted methane against the lifetime of a well (Zavala-araiza *et al.*, 2015). The numbers were generated with a mean of 4.005 Mm^3 and standard deviation of 1.259 Mm^3, where N is the number of samples, which is 10,000 in this case. The preceding calculations for the mean and standard deviation were taking the natural log of the EUR and calculated the mean and standard deviation after that. This was used to produce 10,000 cumulative distribution curves of methane emissions from each dataset and the Monte Carlo simulation plots are shown in Figure 14, 15, 16 and 17 in the results section.

3.4 Skewness analysis and Lorenz curves

An in-depth analysis of the super-emitters in this study is to calculate the degree of skewness for a specific population, whether it is equipment-specific or facility-specific. A certain degree of skewness would indicate a “fat-tail” or “skewed to the right” of the methane emissions’ probabilistic distribution profile. This characteristic indicated a possible existence of the super-emitters as they emit a disproportionate amount of methane. To investigate the skewness for each dataset, the Lorenz curve was utilised.

The Lorenz curve is a graph within the field of economics to represent the inequality of the income distribution for a specific amount of population. The curve presents the cumulative percentage of the total income against the cumulative percentage of the total population. Whilst this curve is used exclusively within economics, it was used to efficiently evaluate the skewness within a methane emission dataset due to the inequality nature of the curve (Zavala-Araiza *et al.*, 2018). The curve would represent the cumulative percentage of the total population of an emission source or reporting segment plotted against the cumulative percentage of methane emissions.

The datasets for both Clearstone and GreenPath were split and categorised according to its emission source and reporting segments. This can be done by using the “dplyr” package. By plotting the Lorenz curves for each emission source and reporting segment, it can be investigated to see which emission source or reporting segment contributes to the skewedness within the dataset. Ultimately, this would be the method to accurately identify the super-emitters in the dataset. However, the datasets from Zavala-Araiza and O'Connell only show methane emissions from gas production sites and not explicitly on the emission source within the facility itself. Nevertheless, the top-down datasets can be used to plot their associated Lorenz curves.

The disadvantage of using cumulative percentage of the total population in the Lorenz curve is that there is a bias for the values of the emitted methane from each emission source or reporting segment due to different natural gas production rates. It would be more desirable to model the cumulative percentage of "Of Throughput" instead of cumulative percentage of the total population of a considered split dataset. Of throughput is defined as methane emissions as a percentage of natural gas production rate, not to be mistaken with the EUR. Unfortunately, the datasets in this study (except Zavala-Araiza's dataset) did not include the natural gas production rates at the time the methane emissions were recorded. This would prevent the bias from appearing within the Lorenz curve analysis and provide better insights on how the super-emitters are characterised based on the relationship between methane emissions and natural gas production rates.

In RStudio, a library called "ineq" was imported as the package specifically provides functions to plot Lorenz curves. The plotted Lorenz curves for the 4 datasets are shown in Figure 18, 19, 20, 21, 22, 23 and 24 in the results section.

After plotting the Lorenz curves, the super-emitters can be further evaluated by investigating the top 5% proportion of each split dataset and plot these against the average absolute emission rate of the original dataset as a simple incentive. By doing this, the trend of heavy-tails can be evaluated if the tails fall under lower or higher emissions on average. This is done by plotting a scatter graph of the top 5% emitters against the calculated average absolute emission rate from each dataset and the plots are shown in Figure 25, 26, 27,28 and 29 in the results section. The visualisation of the plots was enhanced using the "ggplot2" package.

3.5 Economic Assessment: mitigation cost and optimisation model

The incentive for this sub-chapter is to economically assess a variety of mitigating technologies' functionality on reducing the methane emissions in a cost-effective way. The considered type of mitigation technology in this study will exclusively focus on ARPA-E's MONITOR's mitigation technologies. The considered mitigation technologies are a type of LDAR technology as the potential emission reduction by using an LDAR is significantly large relative to other conventional mitigation technologies (Balcombe, Brandon and Hawkes, 2019). The main influence for using MONITOR's mitigation technologies in this study was to promote the technologies' nature for being low-cost and innovative. The program also sparked competitiveness between companies that enrolled onto the program, influenced companies to continuously improve and develop better type LDAR mitigation technologies.

In this study, only two MONITOR technologies from IBM and Aeris Technologies were utilised since both companies provided economic data such as operating expenditure for the economic assessment. Other technologies were not considered in this study as some information were not acquired due to confidentiality reasons. IBM's MONITOR technology is called AIMS (An Intelligent Methane Management Solution) where the technology is a low cost TDLAS (Tunable Diode Laser Absorption Spectroscopy) sensor that provides automated 24/7 continuous monitoring of methane and data storage on IBM's cloud system (IBM, 2019). Aeris' technology is a sensitive mid-infrared laser sensor computed with the atmospheric dispersion model and artificial neural networks for efficient methane detection and quantification (ARPA-E, 2015). The operating expenditures for both technologies are shown in Table 6 below. The considered exchange rate between USD and CAD were $1USD to $1.32CAD, defined as the average yearly exchange rate for the last 5 years (OFX, 2019).

Table 6: Operating expenditures in USD and CAD for the considered mitigation technology costs

Mitigation Technology	OPEX($USD/Year)	OPEX ($CAD/Year)
IBM: AIMS	1526	2000
Aeris Technologies: Natural gas leak detection system	1000	1310

The economic assessment was only applied to the Clearstone and GreenPath datasets as both displayed the source of methane emissions. It would be reasonable to conduct the economic assessment on these datasets; hence the implementation of the technologies would be more specific and precise on tackling the methane emissions.

The methodology for conducting the economic assessment is to reduce the methane emissions to 45% of its original emission value to comply with the Government of Alberta's reduction target. Using the reduced methane emissions value, the mitigation cost for each technology was calculated using the equation below.

$$Mitigation\ Cost = \frac{Total\ technology\ cost - (Reduced\ total\ emission)(Gas\ Price)}{(Reduced\ total\ emission)(GWP)} \quad (3)$$

The mitigation cost ($CAD/tonne CO_2eq) was calculated as a net cost basis using the average 2018 AECO-C natural gas price of $1.48 CAD per GJ. This was later converted into $CAD per tonne of natural gas by multiplying the AECO-C gas price with the natural gas energy content of 53.20GJ/tonne of natural gas. Lower and upper bound of the mitigation costs were evaluated using the lower and upper bound of the GWP100 of 25 and 36 respectively (Umeozor *et al.*, 2019). Costs of IBM and Aeris's mitigation technologies against each emission source are tabulated in Table 13 in the results section.

An optimisation model from CERI was then implemented to minimise the total cost of implementation of the technologies for each emission source whilst attaining the 45% reduction target for both Clearstone and GreenPath dataset. The model is shown below (Umeozor *et al.*, 2019).

$$\min \sum_{i=1}^{N} RE_i \ x \ MC_j \qquad (4)$$

$$s.t. \sum_{i=1}^{N} RE_i \ = 0.45 \ x \ ME_i \qquad (5)$$

$$0 \ \leq \ RE_i \ \leq \ RE_{Max,i} \qquad (6)$$

Where RE_i is the reduced methane emission for emission source *i*, MC_j is the mitigation cost for technology *j*, ME_i is the total methane emission for emission source *i* and $RE_{Max,i}$ is the maximum reduced methane emission for emission source *i.* The optimisation model was formulated in RStudio using the optim function along with the Brent algorithm solver. This solver can only be used in one-dimensional problems; in this economic assessment, the mitigation cost is the variable being optimised so this optimisation model can be considered as one-dimensional. The choice of solvers within the optim function was limited and only Brent was the appropriate solver for one-dimensional problems. The initial guess for the optimisation model for all emission sources was the upper bound IBM/Aeris mitigation cost of -$3.18CAD/tonne CO_2eq from Table 13. The results from the optimisation model for each emission source are shown in Table 14 in the results section.

3.6 Economic Assessment: Sensitivity Analysis via Monte Carlo Simulation

The final step of the economic assessment was to conduct a sensitivity analysis on the AECO-C natural gas price and the mitigation technology cost for IBM/Aeris with the help of Monte Carlo Simulation. The motivation for the sensitivity analysis is to model the fluctuations in the AECO-C natural gas price within 10 years (2018 to 2028) and its effects on mitigation technology cost. The AECO-C base natural gas price from 2018 to 2028 are shown in Table 7 below (Alberta Energy Regulator, 2019).

Table 7: AECO-C Natural Gas Price 2018-2028. Source: (Alberta Energy Regulator, 2019)

Year	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
AECO-C Gas Price ($CAD/GJ)	1.48	1.75	2.19	2.55	2.62	2.76	3.01	3.15	3.35	3.44	3.52

A Monte Carlo simulation was then performed by randomly draw 10,000 estimates of the AECO-C natural gas price from the normal distribution with a mean and standard deviation were calculated as $2.71CAD/GJ and $0.680CAD/GJ respectively. A histogram of the AECO-C natural gas price is shown in Figure 30 in the results section.

Sensitivity analysis of the mitigation cost for each mitigating technology was constructed by executing another Monte Carlo simulation by randomly draw 10,000 estimates of the mitigation cost. The 10,000 random draws were also drawn from the normal distribution with a mean and standard deviation of -$4.88CAD/tonne CO_2eq and $1.51 CAD/tonne CO_2eq for both mitigation technologies. The mean and standard deviation of the mitigation costs were accounted for the different AECO-C natural gas prices in Table 7. The histograms of the IBM and Aeris' mitigating technologies are shown in Figure 31 and 32 in the results section.

Chapter 4: Results and Discussion

4.1 Qualitative and quantitative analysis of probabilistic emissions model

The cumulative distribution curve figures showed that Weibull and Gamma distributions were underestimated relative to the original empirical cumulative distribution curve. Whilst the log-normal and log-logistic distribution showed that the empirical curve and the two distribution curves coincided adequately. It would be probable that Weibull and Gamma distributions are not appropriate distributions in modelling methane emissions. However, to decide on the best fitting curve, the statistical quantitative tests were analysed.

Table 3 shows the maximum log-likelihood of each dataset and its associated distribution curve. The distribution curves were decided based on the values from Table 3 as the curves were produced using maximum log-likelihood estimation. The estimation was used to determine likelihood fit of the empirical cumulative distribution curve with a certain statistical distribution. Maximum likelihood values between the log distributions were all found to be higher than the Weibull and Gamma, so this can be inferred that the Weibull and Gamma are inappropriate distributions to model the considered methane emissions dataset. As both likelihood values for Weibull and Gamma were found to be the lowest values relative to the log distributions. Although this should not be an indication that Weibull and Gamma cannot be used within the methane emissions field due to the bias of the amount of collected methane emissions dataset in this study. Balcombe et al discovered certain natural gas supply chain stages such as the drilling and liquids unloading for manual/automatic plungers can be fitted with Gamma and Weibull distribution respectively (Balcombe, Brandon and Hawkes, 2019).

Table 3 showed that the log-logistic distribution curves were the best distribution curves as it has the highest value out of all the four distributions. Previously in Section 3.2, it was stated that the log-logistic distribution was not selected to represent the statistical distribution due to the nature of the distributions' input values. Log-normal distribution was chosen instead, and it was appropriate as the likelihood values are only slightly different from the log-logistic distribution's values.

Table 4 and 5 shows the goodness-of-fit tests and criteria, which are used to support the likelihood fit further. Akaike's information criterion is a type of goodness-of-fit criteria that evaluates the balance between under and overfitting of the statistical distribution curves. Essentially, it is a criterion test that ranks the quality of the fitting models and the best value for the AIC is the AIC value with the lowest score (de Leeuw and Akaike, 1973)(Snipes and Taylor, 2014). The Bayesian information criterion (BIC) operates the same way as the AIC, but with the difference that the BIC contains a larger penalty term for adding more parameters that could result in an overfitting of a model (Schwarz, 1978). Table 4 and 5 shows that the criterion tests for the log-logistic distributions are the lowest and log-normal distribution's criterion tests were the second lowest. This was consistent with the ranking of the maximum likelihood values.

The final quantitative assessment is to evaluate the goodness-of-fit tests of Cramer von-Mises and Anderson-Darling tests. Cramer von-Mises statistical test is an exclusive test for cumulative distribution functions used to evaluate the distances between the empirical and fitted cumulative distribution curves as a measure of goodness-of-fit. The Anderson-Darling statistical test is a goodness-of-fit test of the data against a specified distribution (Minitab, 2019) (Anderson and Darling, 1952). Both statistical tests must show that they are the lowest out of the chosen distributions, in this case, the log-logistic distribution has the lowest values for both tests. Again, this was consistent with the values of the maximum log-likelihood and the goodness-of-fit criterions.

However, Brandt et al noted that log-normal distributions tend to underestimate the heavy-tail of the distributions as the emission leaks were not classified as categorical dependent variables. Yet, utilising other statistical distributions such as log-logistic distributions, which produce heavier-tails than the log-normal distribution, would also lead to larger uncertainties (Brandt, Heath and Cooley, 2016). Whilst some datasets exhibit specific statistical distributions, there is a bias in this study that the number of populations in the dataset are not identical. There is an incentive that aggregating the data altogether is ideal but it would prove to be challenging based on many factors such as methodology on data collection, the year it was collected, etc.

This section was concluded that the methane emissions and the EUR data in this study were characterised as log-normal distributions even though log-logistic distribution holds the best fit. The results from this section should provide incentives on how to characterise methane emissions using the four distributions to identify super-emitters, as their underlying characteristics are to quantify the right-skewedness of the four distributions. Qualitative and quantitative tests should be utilised to decide the best fit based on the balance between under-fitting and over-fitting of the statistical distribution curves.

4.2 Monte Carlo simulation analysis

The cumulative distribution curves of methane emissions from performing the Monte Carlo simulation specified in Chapter 3.3 are shown below in Figure 14, 15, 16 and 17

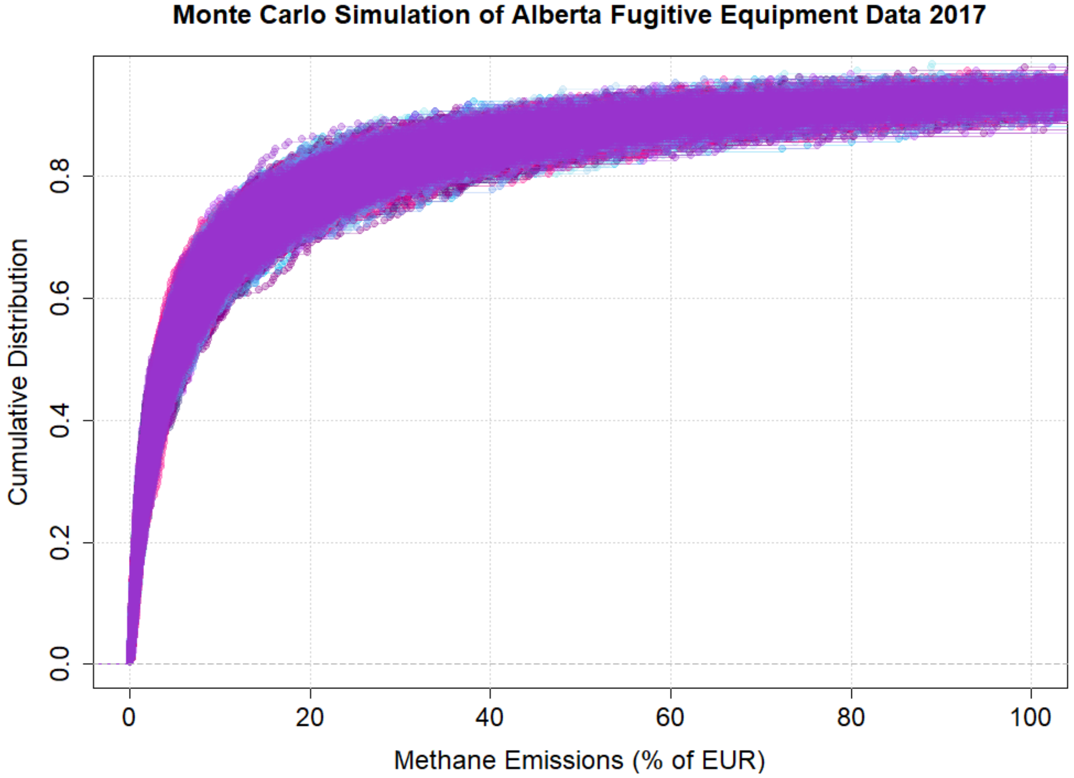

Figure 14: Monte Carlo simulation of methane emissions (% of EUR) for Clearstone dataset

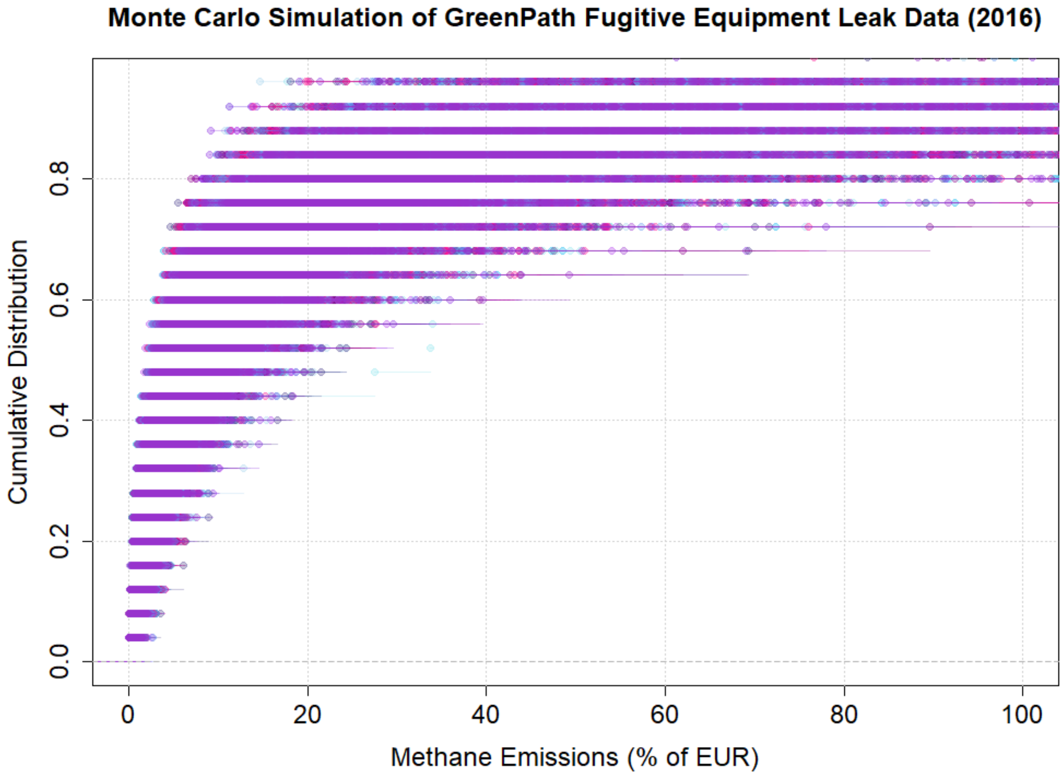

Figure 15: Monte Carlo simulation of methane emissions (% of EUR) of GreenPath dataset

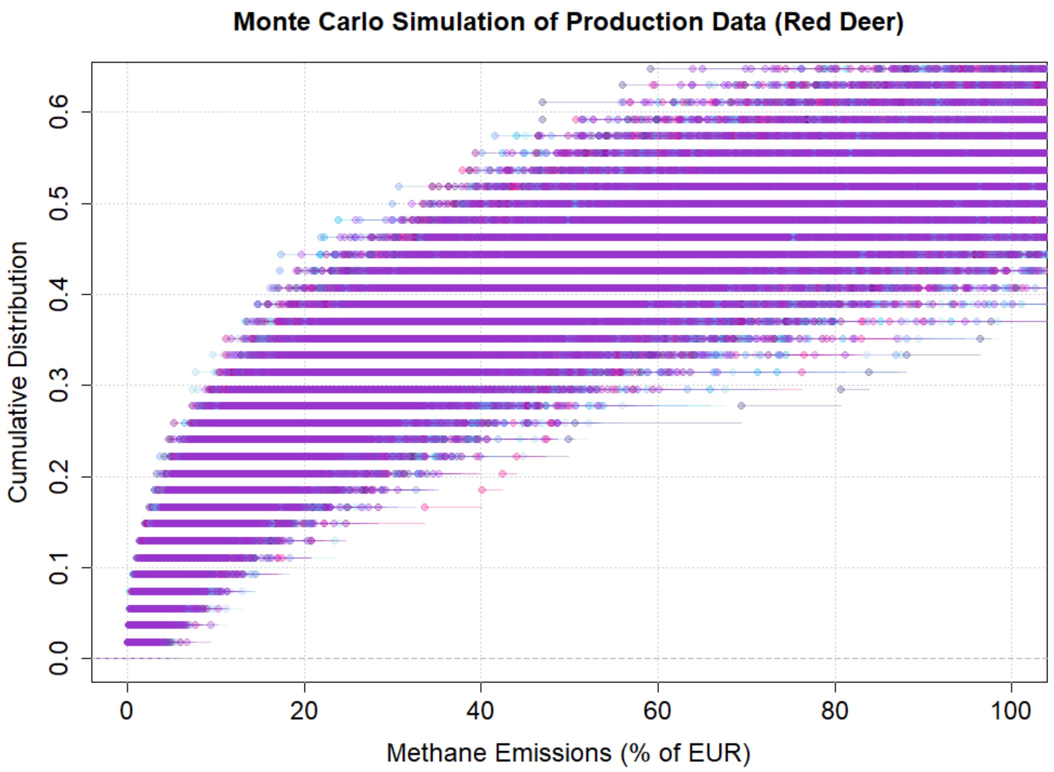

Figure 16: Monte Carlo simulation of methane emissions (% of EUR) of Zavala-Araiza dataset

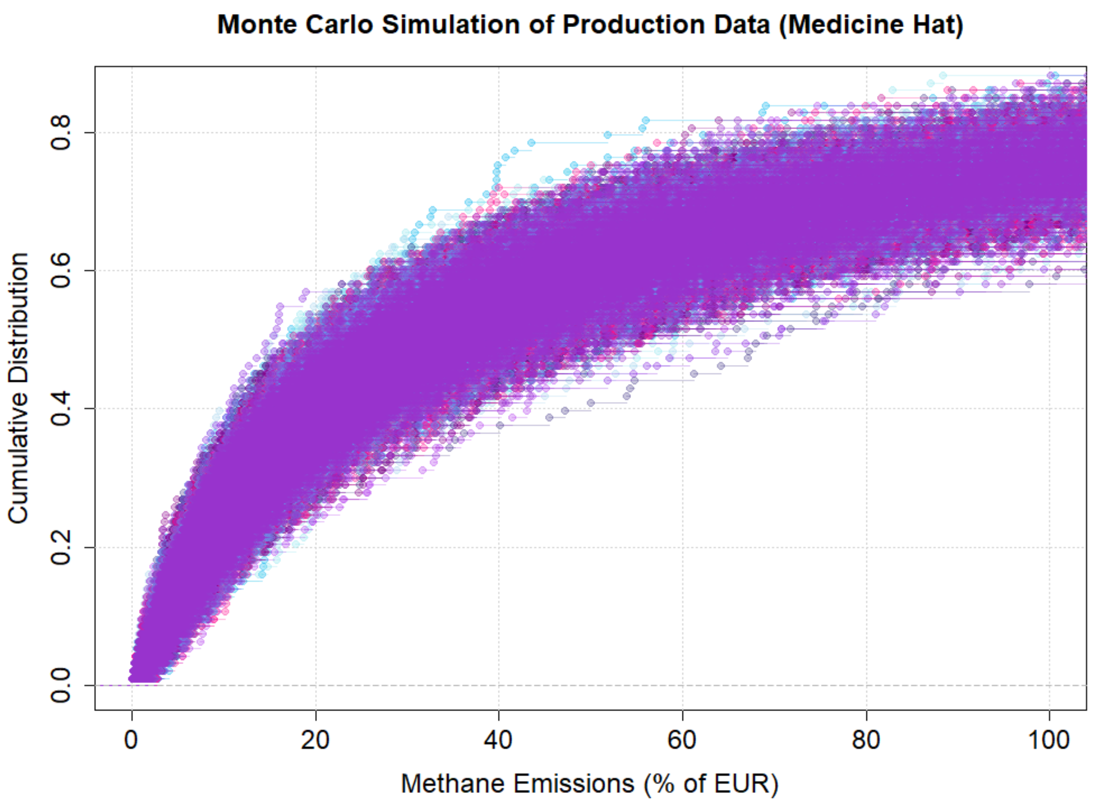

Figure 17: Monte Carlo simulation of methane emissions (% of EUR) for O'Connell dataset

During the Monte Carlo simulation, there were several methane emissions that exceeded substantially from the specified randomly generated EURs. The range for this goes up to 100-3500%. It is extremely unlikely the exceeded scenarios will occur as the methane from the EUR production would be completely extracted by the 100% scenario. Only 100% scenarios were considered in the Monte Carlo simulation analysis.

All Monte Carlo curves could not be distinguished from each other, instead the range of the 5th,25th,50th,75th and 95th percentile emission estimates were calculated for each dataset and this is shown below in Table 8.

Table 8: 5th, 25th, 50th, 75th and 95th percentile ranges for each dataset

Source	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
(Clearstone Engineering Ltd., 2018)	0.0899-0.539%	0.713-2.186%	2.475-7.261%	9.278-32.137%	59.343-100%
(GreenPath Energy Ltd., 2016)	0.0824-3.576%	0.514-9.405%	1.849-33.766%	5.501-94.195%	14.006-100%
(Zavala-Araiza *et al.*, 2018)	0.281-12.284%	6.199-55.337%	30.319-100%	N/A	N/A
(O'Connell *et al.*, 2019)	0.569-6.962%	4.630-22.574%	15.448-70.552%	39.831-100%	N/A
Total	0.0824-12.284%	0.514-55.337%	1.849-100%	5.501-100%	14.006-100%

Table 8 shows that there is tremendous variation within the percentile ranges of the methane emissions, and this was to be expected as all datasets were log-normally distributed. There are discernible differences between the bottom-up and top-down datasets that the top-down datasets showed larger methane emissions distribution in the 5th,25th,50th and 75th percentiles relative to the bottom-up datasets. Certain percentile ranges that displays N/A indicate that emissions exceeded the 100% range and is not included in the table. The curves for the top-down bottom datasets were less steep than the bottom-up datasets, which implied that the ranges of methane emissions were significantly higher than the bottom-up datasets.

As previously stated in Section 2.4, the nature of the top-down methods leads to overestimation of the methane emissions because of the complex difficulty in specifying atmospheric methane concentrations to the emission sources. In addition, as the distance between the source and the airborne measurement device increases, the accuracy in obtaining quality methane emissions data were reduced and leads to overestimation of the methane emissions (Balcombe *et al.*, 2017). Overestimation also occurred as a result of aggregation of methane emissions from other industrial sectors such as emissions from cattle, agricultural sector, etc. Hence, it is inaccurate to compare the percentile ranges between the bottom-up method and top-down datasets.

Both bottom-up datasets showed scenarios of the remaining 10-20% of total population emit emissions between 20-100%. Whilst other scenarios were the 30-50% of the remaining total population emit emissions between 20-100%. Majority of the methane emission sources from the two bottom-up datasets are fugitive, leaked emissions and indicate that within the stated scenarios, there is a chance that the extreme variation in the range of 20-100% methane emissions, especially emissions of more than 70% were contributed by the super-emitters. It is interesting to note that the percentile range of emission values in the GreenPath dataset is notably broader than the Clearstone dataset. This was partly due to the large differences of the sample population for both datasets, where there are 25 samples in GreenPath and 195 samples in Clearstone. In addition, there was only one methane emission value of 1526 m^3/day that immensely skewed the percentile range, since the range of the methane emissions of the GreenPath dataset are 0.50 - 25 m^3/day. It would be reasonable to state that some of the remaining 10-20% of total population emitting a range of 20-100% methane emissions were contributed from 1526m^3/day alone and could be classified as a super-emitter.

For the top-down datasets, the included scenarios were 50-60% and 75-85% of remaining total population to emit 20-100% methane emissions. The underlying emission source and its type could not be discussed as the emissions were collected airborne. However, the Monte Carlo results indicated that the top-down scenarios were more skewed relative to the bottom-up datasets. There is a possible chance this was contributed by the super-emitters. Again, the uncertainties associated with the top-down methodology would be hard to corroborate that super-emitters were the major contributions. O'Connell et al's gaussian dispersion model methodology were efficient enough to identify high-emitting sites although the results were overestimated. Fortunately, the methane emissions data were consistent with GreenPath's data but with higher orders of magnitude (O'Connell *et al.*, 2019). With this, it can be corroborated that super-emitters are the main contributors to the skewed emissions in O'Connell's dataset.

It was worth noting that a low EUR contributed to such high methane emissions and it was reasonable to state that the EUR is inversely proportional to the methane emissions. This is because throughout the lifetime of a well, methane emissions were bounded by the material and energy consumption to conduct extraction activities such as site preparation, drilling, etc. Should the production of natural gas failed, the impacts towards methane emissions for having a low EUR would be extremely high per m^3 of gas output (Stamford and Azapagic, 2014). It can be concluded that the EUR is one of the key factors in quantifying the hazardous super-emitters into existence. With 10,000 scenarios, it was statistically robust enough to identify the wide variation of the contribution by the super-emitters, concluding that the super-emitter problem is extremely prevalent and exists within the natural gas supply chain.

Referring to GreenPath's super-emitter, there might be more than one super-emitter than skewed the percentile ranges for both bottom-up dataset, whilst top-down datasets must deal with uncertainties before corroborating judgements. For both bottom-up and top-down emitters, it is imperative to conduct an efficient analysis on the range of super-emitters and its effects on skewing the distribution. This will be done by discussing the results of the Lorenz curves and the top 5% emitter scatter graphs, which will be further discussed in Section 4.3.

There could be other factors that would contribute to the existence of super-emitters other than the EUR. It is a fact that abnormal process conditions such as leakages or improper operation of equipment to exhibit super-emitter behaviours, however, it would be difficult to "predict" when these abnormal process conditions will occur (Brandt, Heath and Cooley, 2016). Proper methane regulations by implementing continuous monitoring devices will prevent the frequency of the occurrence of super-emitters and ultimately reduce the impact of the EUR, ideally at the instantaneous time the methane was emitted. This solution will be further discussed in Section 4.4.

4.3 Skewness and Lorenz curves analysis

The Lorenz curves for all datasets are shown below in Figure 18, 19, 20, 21, 22, 23 and 24.

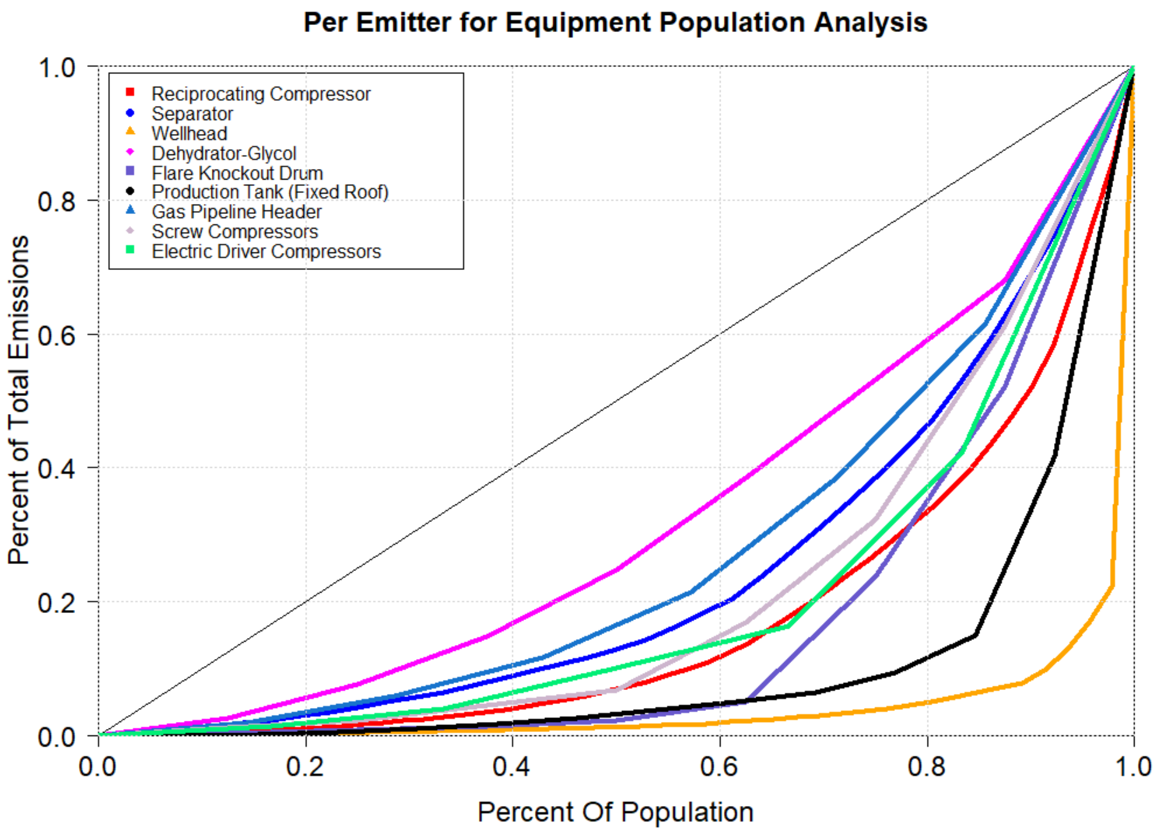

Figure 18: Lorenz curve of "per emitter" analysis for Clearstone's emission source

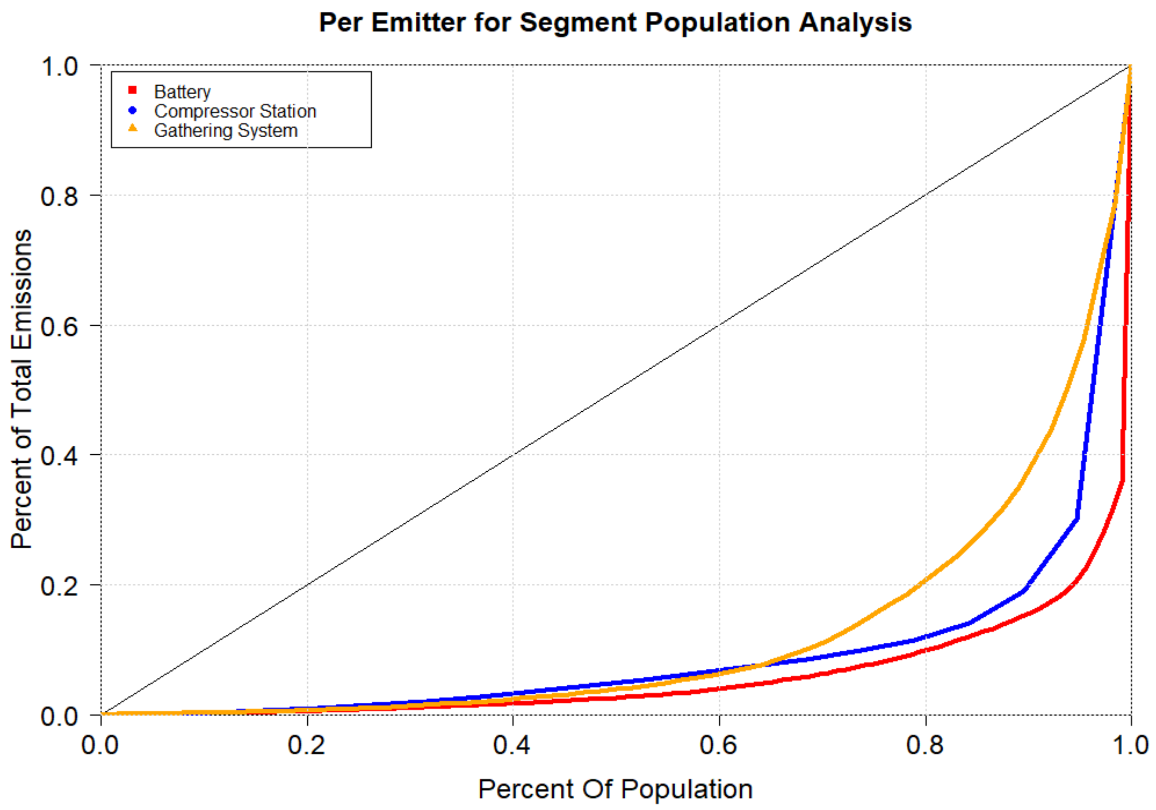

Figure 19: Lorenz curve of "per emitter" analysis for Clearstone's reporting segment

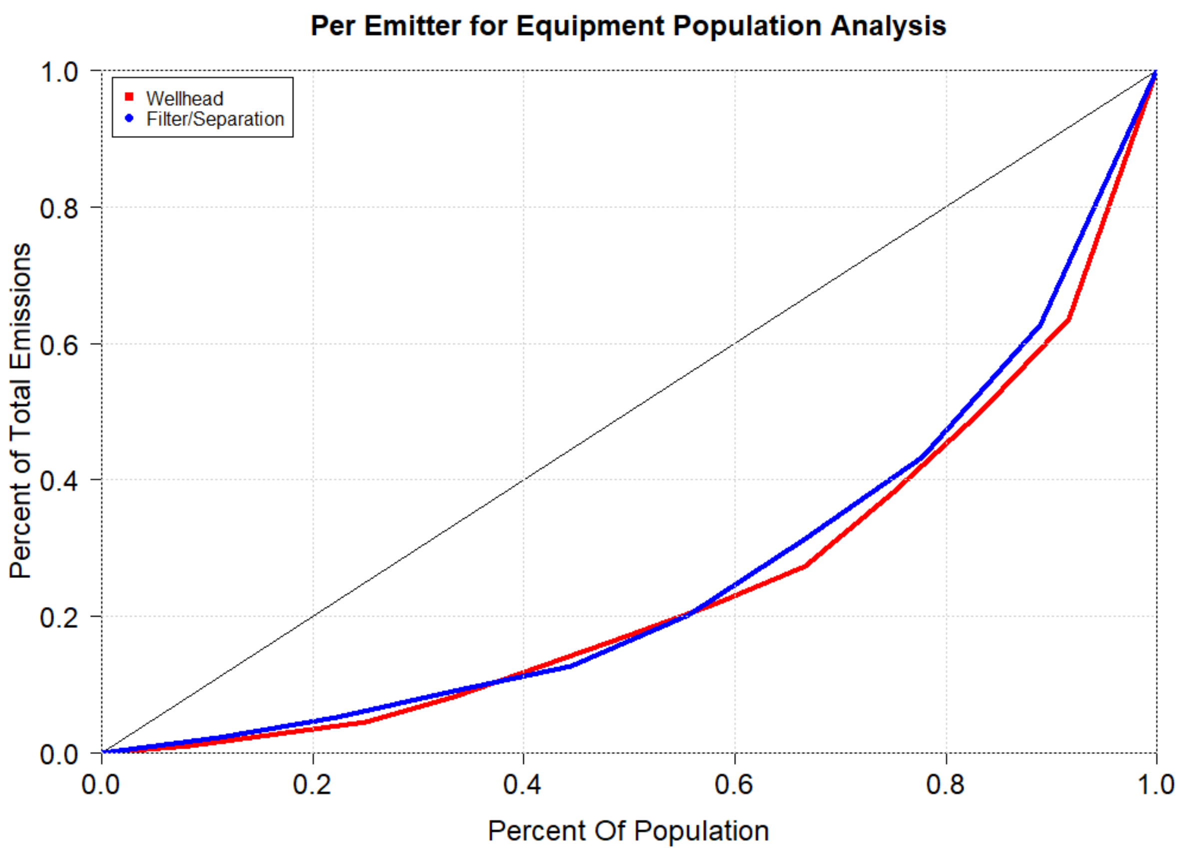

Figure 20: Lorenz curve of "per emitter" analysis for GreenPath's emission source

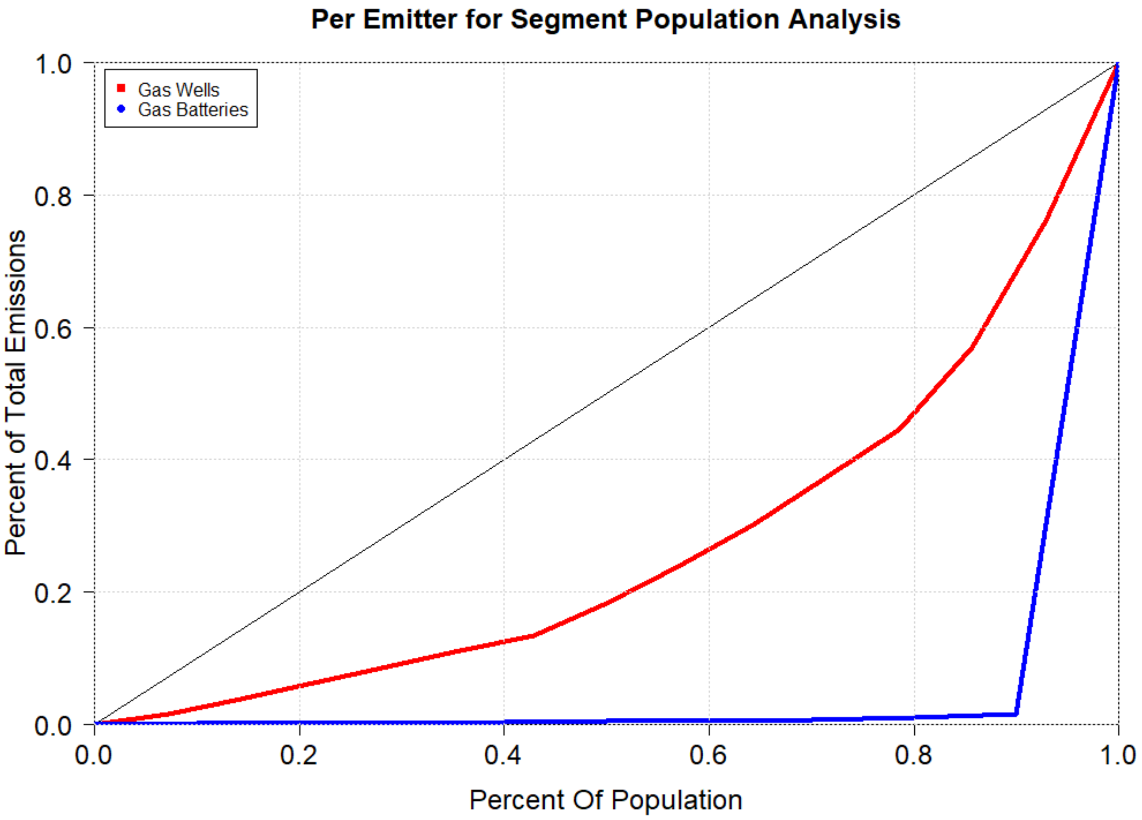

Figure 21: Lorenz curve of "per emitter" analysis for GreenPath's reporting segment

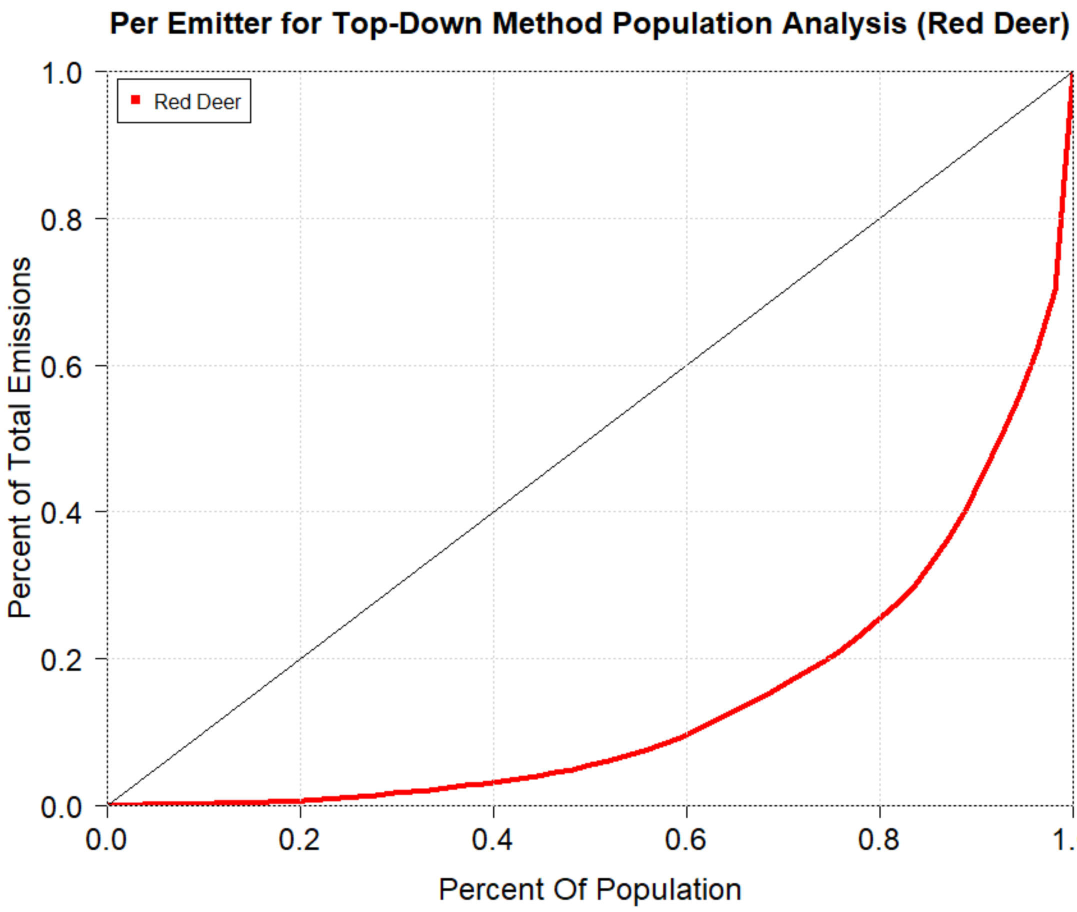

Figure 22: Lorenz curve of "per emitter" analysis for Zavala-Araiza dataset

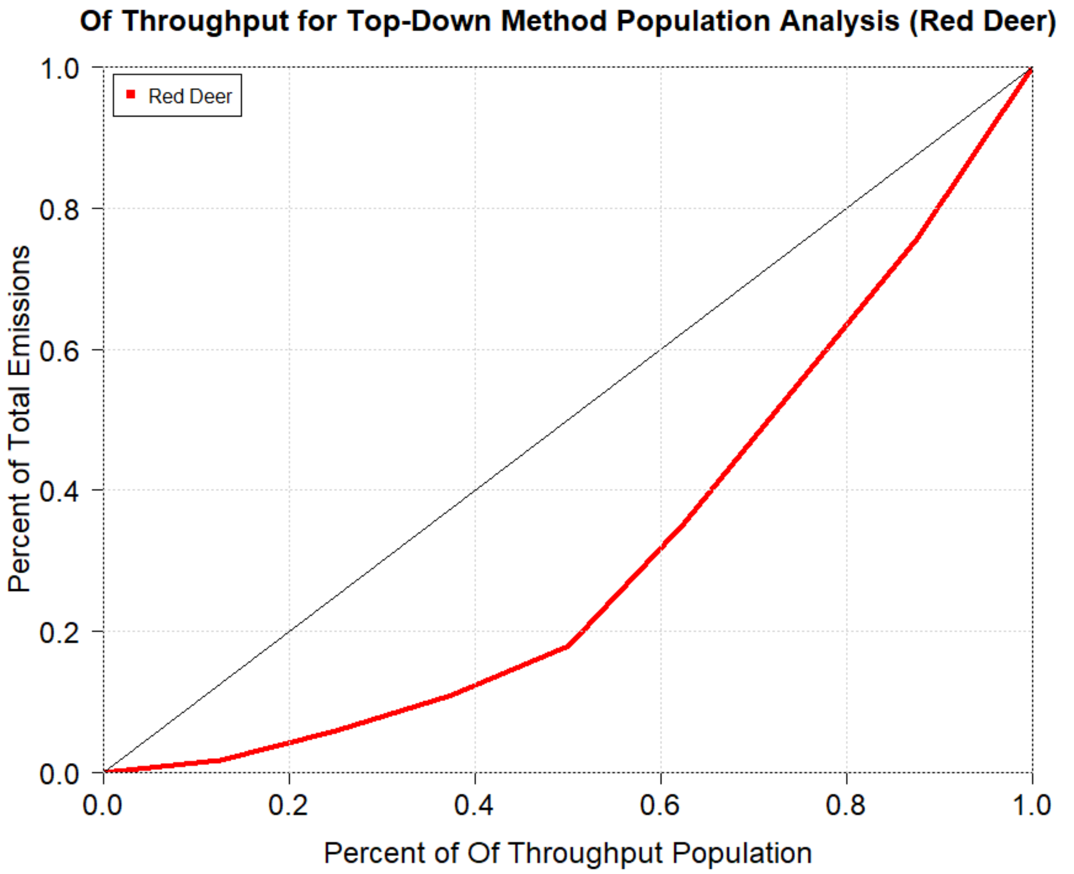

Figure 23: Lorenz curve of "of throughput" analysis for Zavala-Araiza dataset

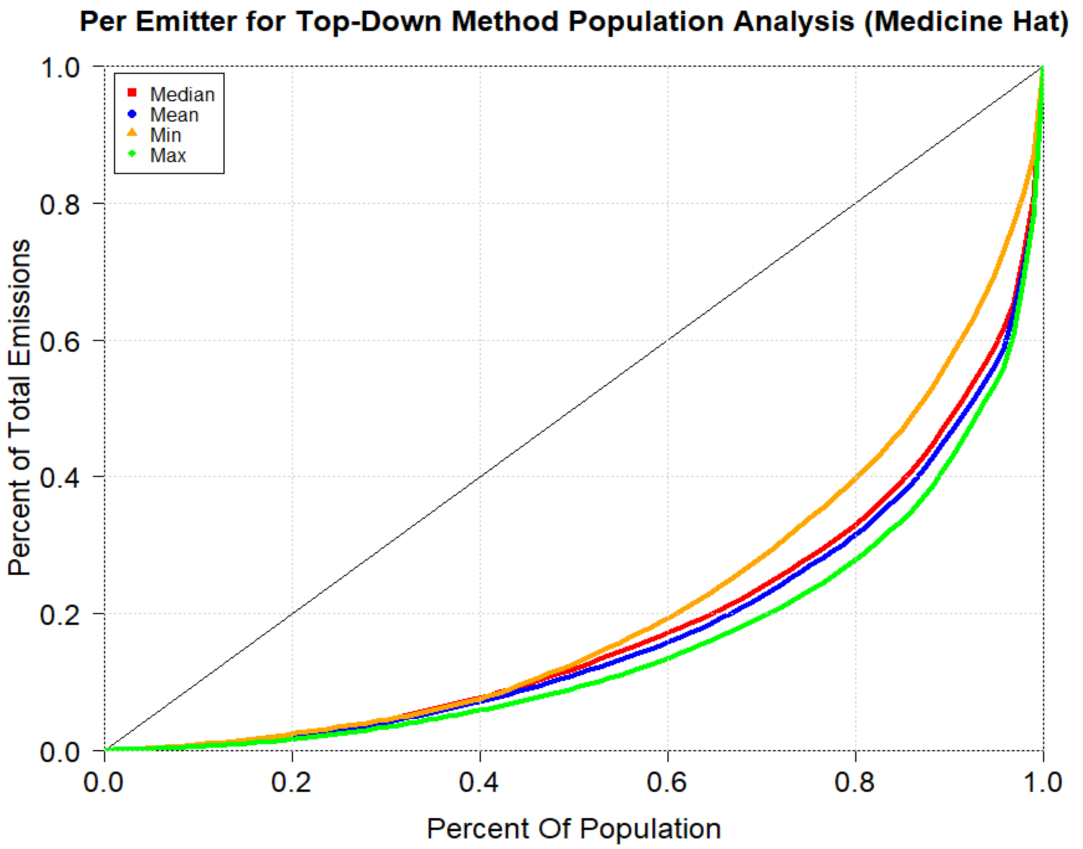

Figure 24: Lorenz curve of "per emitter" analysis for O'Connell dataset

A rule of thumb in analysing the Lorenz curves is to utilise the 80/20 rule (the Pareto principle) to investigate the contribution by the super-emitters. The 80/20 rule applied to the methane emissions context would imply that 20% of total methane emissions were from 80% of the total population. With the other 20% of total population emits 80% of the total methane emissions. This rule was used by GreenPath as an incentive to evaluate the super emitting sources within the field survey in order to propose the LDAR's efficiency in mitigating large emission sources (GreenPath Energy Ltd., 2017)

Starting off with the reporting segment population figures, Figure 19 showed that 80% of the total population for reporting segments emit 20% or less of total methane emissions. For compressor stations and gas batteries, it shows approximately 10-15% of total methane emissions being emitted by 80% of total population. Whilst the gas gathering system curve shows 20% of total methane emissions, which is relatively small to the two reporting segments. From the curves, it can be inferred that for compressor stations and gas batteries, the remaining 20% of total population emitted 85-90% of total methane emissions whereas gas gathering systems emitted 80% of total methane emissions. This indicates that the data within the reporting segments are extremely skewed. With this extremum of skewness, there is a high possible chance that it is contribution from the super-emitters within the reporting facilities, although this does not pinpoint the emission source. Figure 18 provides the solution to the previous statement as it reports the Lorenz curves for the equipment-specific sources.

The curves from Figure 18 show a mix of degree of skewness, some curves show little skew and some curves show more skew. Equipment that shows little skewness to the data are screw compressors, separators, gas pipeline headers and dehydrator-glycols. The remaining equipment that shows a certain degree of skew are electric driver compressors, flare knockout drums, reciprocating compressors, production tanks and wellheads. The difference in the degree of skew were likely due to the differences in population amount for each considered equipment source. Whilst this presented some bias, the contribution towards the skewness within the data is not significant. However, the significant contribution would be the interval between the cumulative methane emissions at 80% and 20% of total population was not constant for each equipment source.

Figure 18 shows the most significant skewedness belongs to the reciprocating compressors, production tanks and wellheads population. This is because 20% of the total population for the stated emission sources shows 65%, 90% and 95% of total methane emissions. To see the significant contribution of the equipment sources to the reporting segment, a table of the top 5% emitters for the reporting segments is shown below in Table 9.

Table 9: Top 5% emitter(s) for Clearstone's reporting segments

Reporting Segment	Emission Type	Emission Source	Emission Volume (m^3/day)
Gas Gathering System	Leak	Reciprocating Compressor	80.900
Gas Gathering System	Unlit Flare	Flare Knockout Drum	84.214
Gas Gathering System	Leak	Production Tank (Fixed Roof)	165.726
Gas Gathering System	Leak	Reciprocating Compressor	57.262
Compressor Station	Vent	Production Tank (Fixed Roof)	360.216
Gas Battery	Vent	Production Tank (Fixed Roof)	34.725
Gas Battery	Leak	Wellhead	63.272
Gas Battery	Leak	Wellhead	75.828
Gas Battery	Leak	Wellhead	41.933
Gas Battery	Leak	Wellhead	1126.646
Gas Battery	Leak	Reciprocating Compressor	52.980

Table 9 shows the top 5% emitters within the reporting segment population, which are classified to be the super-emitters in this dataset for the reporting segment. Wellheads were the most significant contributor to the extreme skewedness within the gas battery segments since there was a leaked emission of 1126.646 m^3/day. For compressor stations, there was only one super-emitter in contrast to gas gathering system's four super-emitters. Yet, compressor stations showed more skewedness than the gas gathering systems segment. This was due to the significantly large magnitude of emission volume from the compressor station segment. Hence, the skewness contribution from the emission sources to the reporting segments' Lorenz curves are consistent in Figure 18 and 19.

Figure 20 and 21 show the Lorenz curves for the emission source and reporting segments in the GreenPath dataset. Starting from the reporting segment Lorenz curves, there is a stark difference in the skewedness for gas wells and gas batteries. Gas wells showed that the remaining 20% of total population emitted 50% of the remaining total emissions. However, gas batteries showed that approximately 98% of the total emissions were emitted by the remaining 10% of total population. Surprisingly, the Lorenz curves for the emission sources showed similar skewedness, where both wellheads and separators showed 20% of the remaining population emitted 50% of total emissions. It shows that the Lorenz curves for the emission source and reporting segments are slightly inconsistent. To see the emission sources' contribution, Table 10 shows the top 5% emitters for the reporting segment.

Table 10: Top 5% emitter(s) for GreenPath's reporting segments

Reporting Segment	Emission Type	Emission Source	Emission Volume (m^3/day)
Gas Well	Leak	Filter/Separation (Separator)	25.427
Gas Batteries	Leak	Tankage	1525.608

From Table 10, the extreme large skewedness contribution from gas batteries segment was from the tankage emission source of 1525.608 m^3/day. The magnitude for this emission volume was extremely high relative to all the emission sources in the GreenPath dataset, as the range of the emission sources were 0.50 - 25 m^3/day. In addition, the intervals between the emission volumes for both reporting segment and emission source are narrow compared to the Clearstone's emission volume range. Hence, GreenPath's dataset does not show as much skewness as expected. The inconsistency between the emission source and the reporting segment's Lorenz curves was the exclusion of the tankage emission source. However, the inclusion of the tankage emission source's Lorenz curve would provide an even more inconsistency as there is only one tankage emission source in the GreenPath dataset. Applying the Lorenz curve to only one data would create an ambiguous straight line.

Figure 20 shows similar skewness at the remaining 20% of total population, due to the similar range of values exhibited by both emission sources. Table 11 and 12 shows the dataset of wellheads and separators sources that were used to construct the Lorenz curves.

Table 11: Dataset for GreenPath's separator. Row in bold is the top 5% emitter of the dataset

Reporting Segment	Emission Type	Emission Source	Emission Volume (m^3/day)
Gas Well	Leak	Filter/Separation (Separator)	1.526
Gas Battery	Leak	Filter/Separation (Separator)	2.034
Gas Well	Leak	Filter/Separation (Separator)	2.543
Gas Well	Leak	Filter/Separation (Separator)	2.543
Gas Well	Leak	Filter/Separation (Separator)	5.085
Gas Well	Leak	Filter/Separation (Separator)	7.628
Gas Battery	Leak	Filter/Separation (Separator)	8.137
Gas Well	Leak	Filter/Separation (Separator)	13.222
Gas Well	**Leak**	**Filter/Separation (Separator)**	**25.427**

Table 12: Dataset for GreenPath's wellheads. Row in bold is the top 5% emitter of the dataset

Reporting Segment	Emission Type	Emission Source	Emission Volume (m^3/day)
Gas Battery	Leak	Wellhead	0.509
Gas Battery	Leak	Wellhead	1.017
Gas Battery	Leak	Wellhead	1.017
Gas Battery	Leak	Wellhead	2.034
Gas Well	Leak	Wellhead	2.543
Gas Well	Leak	Wellhead	2.543
Gas Well	Leak	Wellhead	2.543
Gas Battery	Leak	Wellhead	3.051
Gas Well	Leak	Wellhead	6.102
Gas Well	Leak	Wellhead	6.611
Gas Well	Leak	Wellhead	7.628
Gas Well	**Leak**	**Wellhead**	**20.341**

The bolded values in both tables are the top 5% emitters in both emission source. Combining all emission volumes except the top 5% emitters would give 42.718 and 35.598 m^3/day respectively from each table. Since the magnitude of the combined emission volumes are higher than the emission volume exhibited by the top 5% emitters itself, the skewness contribution of the top 5% emitters are not as significant. Nevertheless, the mix degree of skew within the GreenPath dataset still indicated the contribution was provided by the super-emitters, but at a lower magnitude of emission volume.

Figure 22 shows the remaining 20% of the total population emit approximately 75% of the total emissions, which is in line with Zavala-Araiza's result (Zavala-Araiza *et al.*, 2018). Again, this indicates that the contribution was from the super-emitting production sites. However, the source of the super-emitters from Zavala-Araiza was unknown as the methane emissions were collected airborne. But inferences from the Lorenz curve can be made that there is a significant degree of skewness within the methane emissions volume.

Since the methane emissions volume vary based on different gas production rates, figure 22 presented a bias that there was no consideration of the gas production rate. Some high-emitting sites might not be emitting as much methane as they produce gas output. Hence, there is an incentive to evaluate the degree of skew further by modelling methane emissions as "of throughput" basis. Figure 23 shows that the initial 50% of the total "of throughput" population emit approximately 15-18% of total emissions, with the remaining 50% emit approximately 82-85% of total emissions.

Comparing both figures, the "of throughput" figure showed a reduced skewness contribution in contrast to the "per emitter" figure. It can be inferred that using the "per emitter" analysis population, it underestimated the contribution of the super-emitters without accounting for the gas production rate. Using the similar reasoning with the EUR in Section 4.2, it can be corroborated that failing to produce gas at a high output would impact the significance of the methane emissions to be classified as the super-emitters. In addition, Zavala-Araiza et al also stated that the proportional gas loss rate were heavily influenced by the low production gas rates (Zavala-Araiza *et al.*, 2018). Fortunately, this can be inferred from the graph that the super-emitters were contributing more significantly than expected. By comparing the emissions with the instantaneous gas production rate, it is an incentive that should be implemented to model super-emitters efficiently.

Figure 24 shows the minimum and maximum Lorenz curves along with the mean and median curves. The minimum Lorenz curve shows that the remaining 20% of total population emits 60% of total emissions whilst the maximum Lorenz curve shows 75% of total emissions emitted by the remaining 20% of total population. Both mean and median curves emit approximately 70% of total emissions by the remaining 20% of total population. Both Figure 23 and Figure 24 shows similarities in terms of degree of skewness so the same inferences from Figure 23 can be corroborated with identification of super-emitters in the top-down methodology.

All datasets show a certain degree of skewness and to further evaluate how much contribution the chosen super-emitters skewed the data; the top 5% emitters scatter graphs will be evaluated. This is shown in Figure 25, 26, 27, 28 and 29. Input data for these figures are in Table 21, 22 and 23 in the Appendix section.

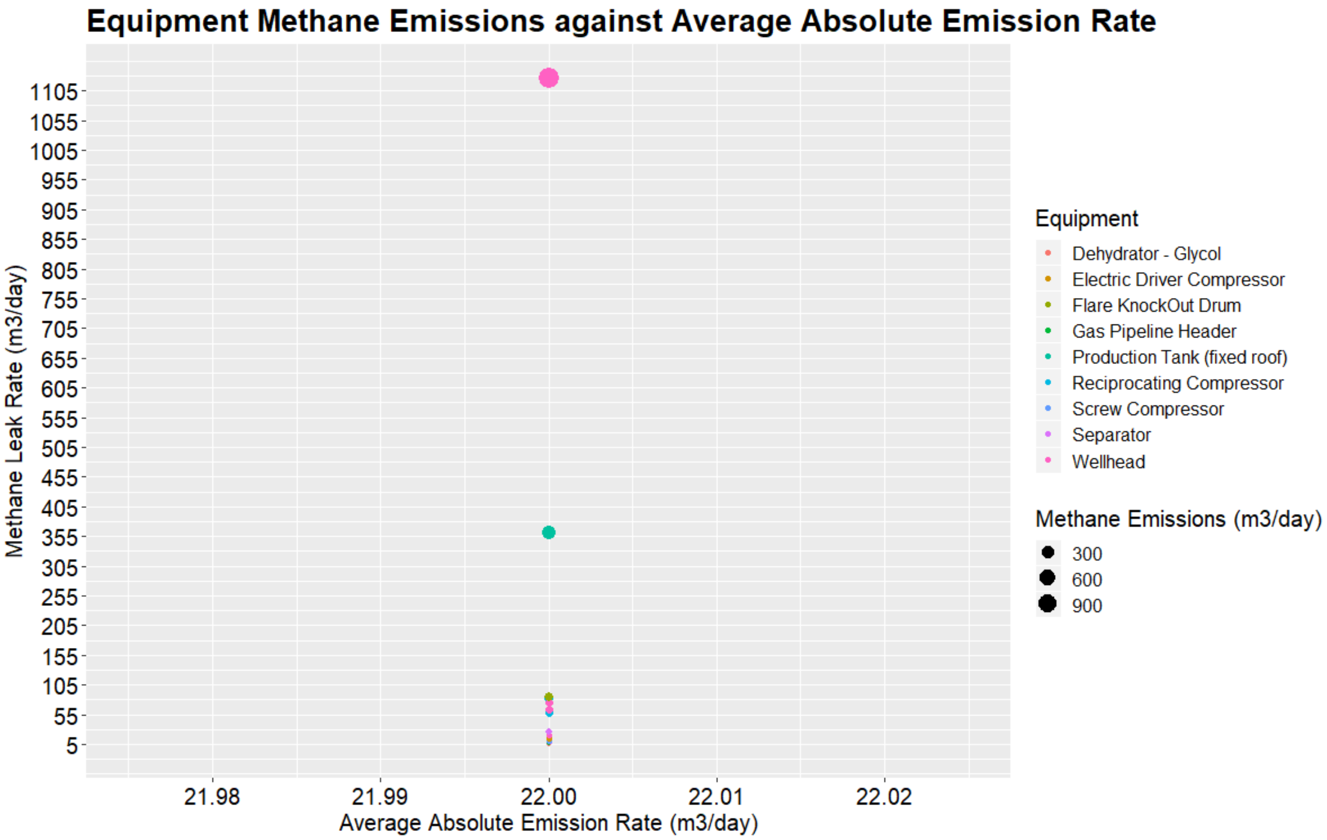

Figure 25: Top 5% emitters from combined emission source population of Clearstone and GreenPath

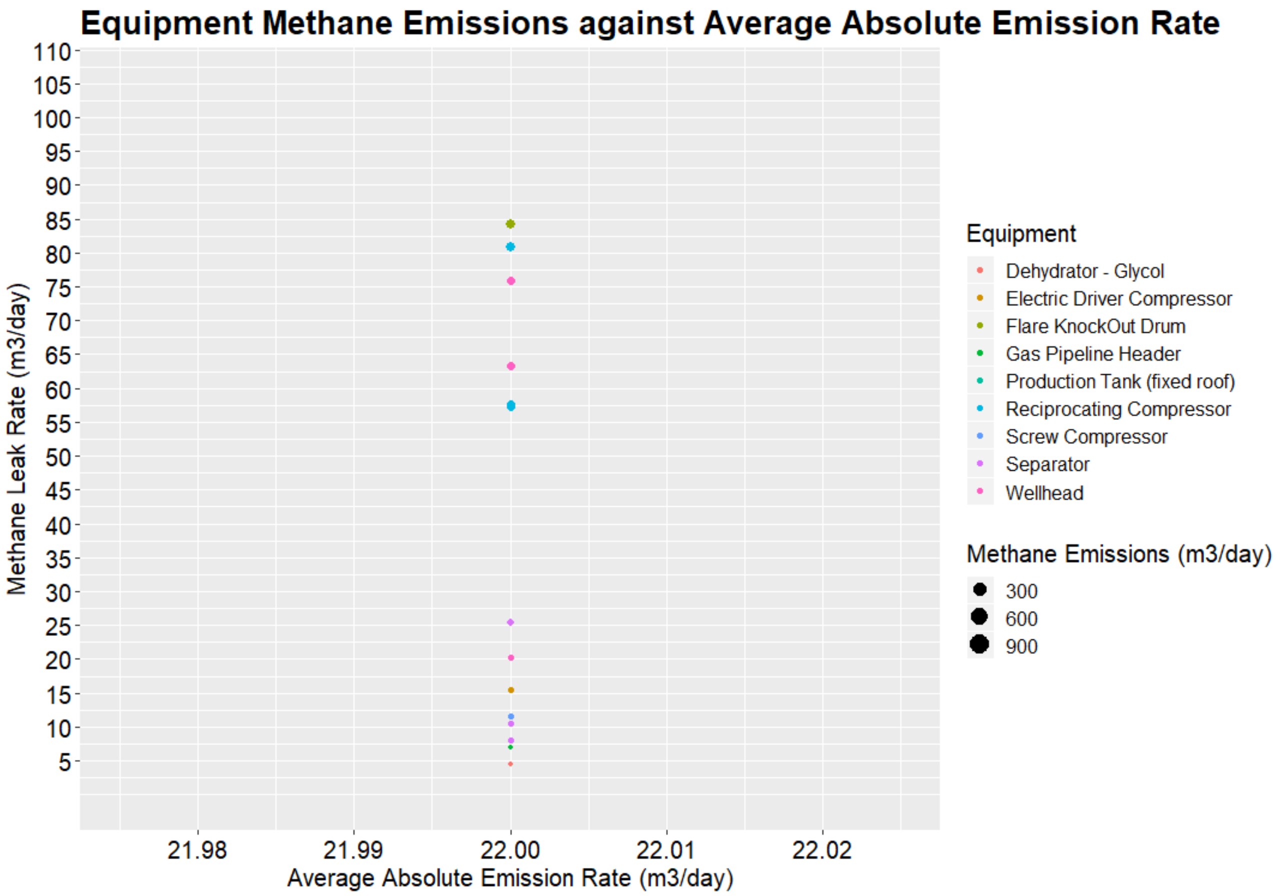

Figure 26: Top 5% emitters 5- 85m^3/day range from combined emission sources of Clearstone and GreenPath

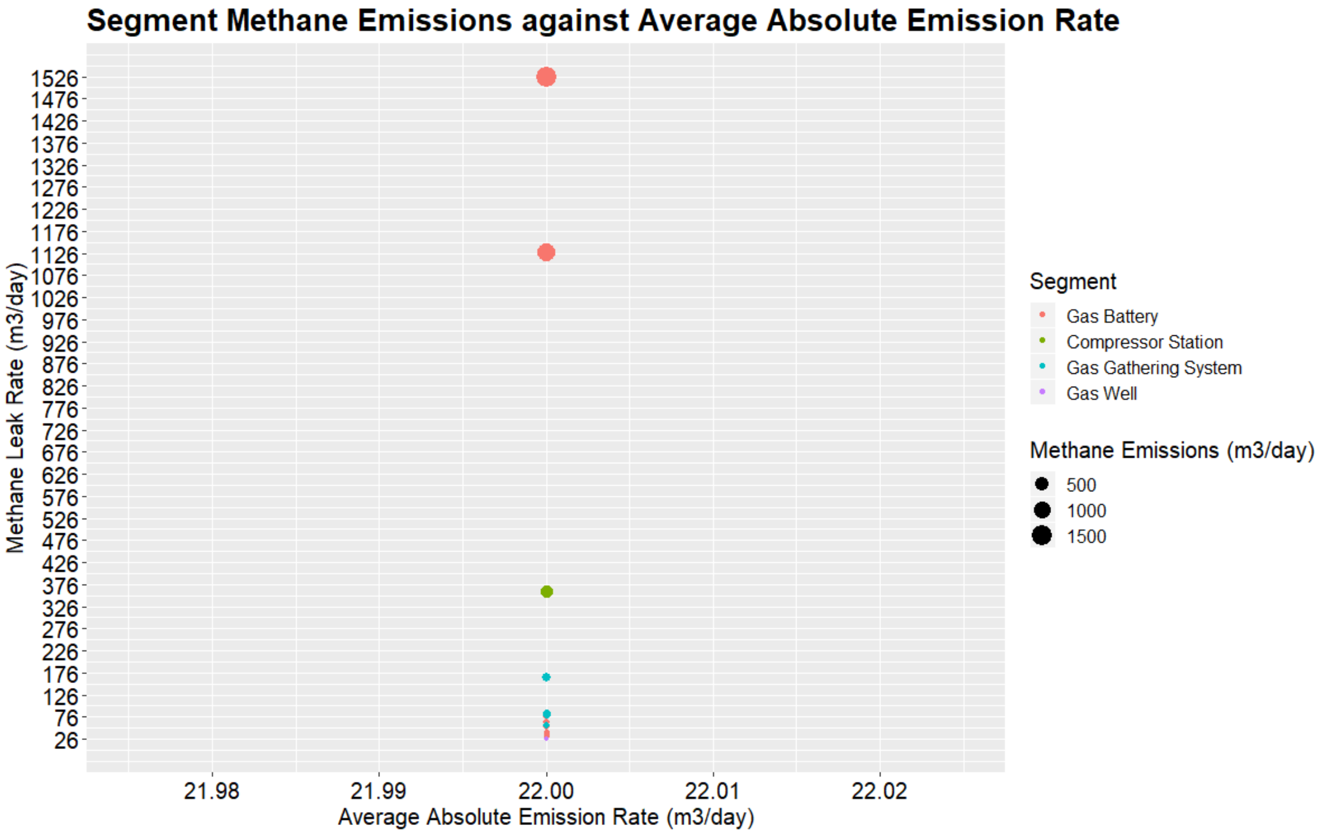

Figure 27: Top 5% emitters range from combined reporting segment population of Clearstone and GreenPath

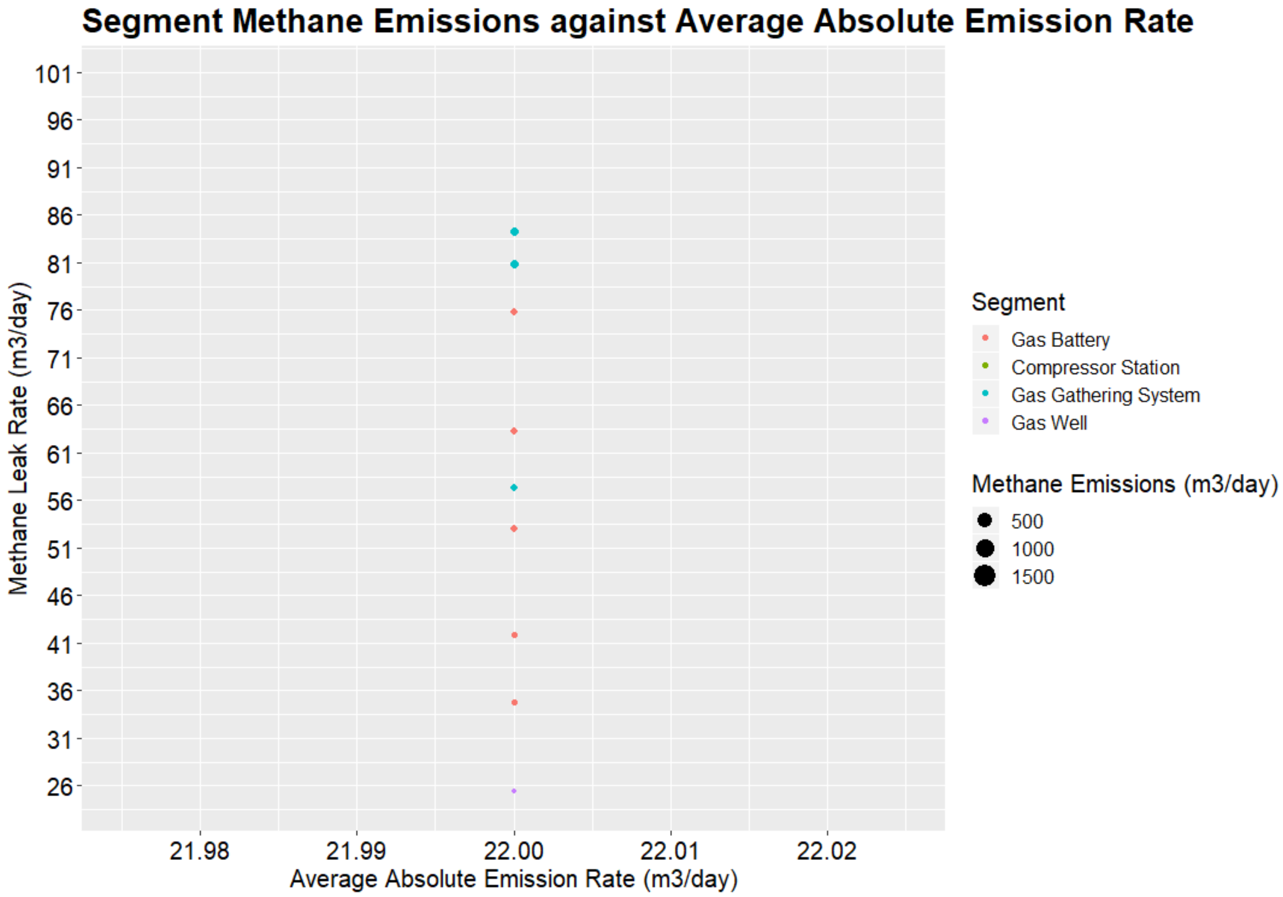

Figure 28: Top 5% emitters range 25 – 86 m^3/day range from combined reporting segments of Clearstone and GreenPath

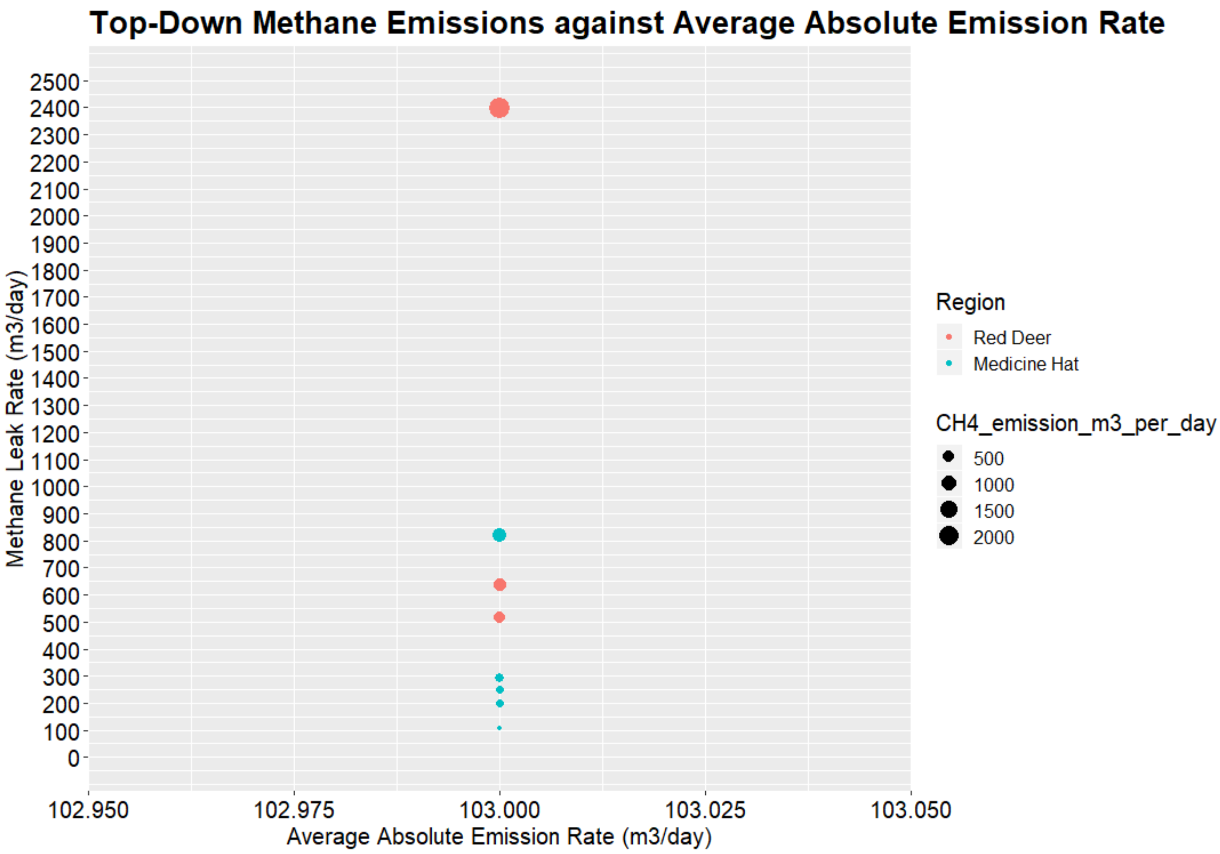

Figure 29: Top 5% emitters from top-down datasets of Zavala-Araiza et al (Red Deer) and O'Connell et al (Medicine Hat)

Starting from Figure 25 and 26 the top 5% emitters are in the range of 5– 1127m^3/day. Typical super-emitting emissions are in the range of 5 – 85m^3/day relative to the average absolute emission rate of 22m^3/day. Evaluating at the range between 25 - 85m^3/day, it seemed that the equipment populations were quite significantly skewed without even accounting for the two highest emitters in Figure 25. With reciprocating compressors and wellheads being one of the large contributors as there were more population than others. Even with the large skewness contribution, there were some top 5% emitters below the average absolute emission rate. Although the combined range of those top 5% emitters were still far lower than the combined range of the top 5% emitters that exceeded the average absolute emission rate. The skewedness contribution from the lower top 5% emitters almost had no effect. However, the trend from both Figure 25 and 26 showed that the heavy-tail characteristics of the equipment population emit methane volume less than 100 m^3/day on average, which is relatively small than some of the extreme top 5% emitters.

Figure 27 and 28 shows the top 5% emitters' range is between 25 - 1525 m^3/day. Again, the heavy-tail characteristics showed that it contains methane emissions smaller than 100 m^3/day on average. For the reporting segment population, all top 5% emitters' emissions exceeded the average absolute emission rate, with 3 methane emission volumes exceeded 100 m^3/day. Even with the average top 5% emitters' emission volume were less than 100 m^3/day, majority of the skewness contribution were from the top 5% emitters with emission volume of more than 165 m^3/day in the gas battery and compressor station reporting segment. For both bottom-up dataset figures, it was indicated that the average emissions emitted by the top 5% emitters were relatively small. However, these emissions did not provide as much skewness as expected, but the biggest skewness contribution would be the top 5% emitters that emit at least more than 100 m^3/day.

Figure 29 shows the range is between approximately 105 - 2400 m^3/day against the average absolute emission rate of 103 m^3/day within the top-down dataset. The heavy-tail characteristics exhibited lower methane emissions on average, similar to the bottom-up dataset's top 5% emitters' average emissions. Where these lower methane emissions are approximately in the range of 105 – 300 m^3/day from the Medicine Hat region. However, the magnitude of the emissions in the top-down dataset were significantly higher than the bottom-up dataset. It is evident that the large magnitude of the emission volumes with more than 100m^3/day was the major contribution to the skewness of the Lorenz curves from Figure 19 and 21. It is interesting to infer that Red Deer emits an extremely high amount of methane relative to the Medicine Hat region. Red Deer is known for inhabiting relatively old gas wells, which would have a significant contribution in the amount of gas production rate. As the age of the gas wells increase, the gas production would decline at a later age in its well lifetime. With low gas production volume, it would highly quantify the number of super-emitters into existence (Guo *et al.*, 2017).

The 5 figures showed that the top 5% emitters skewed the methane emission distribution on a relatively small magnitude of methane emissions volume. However, the skewness contribution from these average emissions emitted by the top 5% emitters was not as significant as expected. In fact, typical emission volumes of more than 100 m^3/day immensely skewed the methane emissions and there would be at least one emitter to come into existence. In addition, the majority of these top 5% emitters exceeded the average absolute emission rate exhibited by each combined bottom-up and top-down dataset. It shows how significant these top 5% emitter can contribute to the total greenhouse gas emissions.

4.4 Economic assessment analysis

The mitigation costs for both IBM and Aeris' technology applied to all bottom-up data are shown in Table 13.

Table 13: Lower and upper bound mitigation technology costs

Emission Source	LB IBM Mitigation Cost ($CAD/tonne CO2eq)	UB IBM Mitigation Cost ($CAD/tonne CO2eq)	LB Aeris Mitigation Cost ($CAD/tonne CO2eq)	UB Aeris Mitigation Cost ($CAD/tonne CO2eq)
Clearstone: Reciprocating Compressors	-2.187	-3.149	-2.187	-3.149
Clearstone: Separator	-2.187	-3.149	-2.187	-3.149
Clearstone: Wellheads	-2.187	-3.149	-2.187	-3.149
Clearstone: Dehydrator-Glycol	-2.187	-3.149	-2.187	-3.149
Clearstone: Flare Knockout Drum	-2.187	-3.149	-2.187	-3.149
Clearstone: Gas Pipeline Header	-2.187	-3.149	-2.187	-3.149
Clearstone: Production Tank (Fixed Roof)	-2.187	-3.149	-2.187	-3.149
Clearstone: Screw Compressor	-2.187	-3.149	-2.187	-3.149
Clearstone:	-2.187	-3.149	-2.187	-3.149

Electric Driver Compressor				
GreenPath: Wellheads	-2.187	-3.149	-2.187	-3.149
GreenPath: Separators	-2.187	-3.149	-2.187	-3.149

Table 13 shows that the average lower bound and upper bound mitigation costs are -2.187 and -\$3.149CAD/tonne CO_2eq for both technologies. Negative mitigation costs indicate that the cost will be returned as profit. However, the calculation of the mitigation cost was extremely site-specific and therefore subject to variability and uncertainties. This methodology was similar to the mitigation cost calculated by (ICF International, 2014). This was in line with the mitigation cost of the implemented LDAR from Figure 5. The reduction percentage applied by the ICF for the implemented LDAR was 60% (with quarterly checks) whilst in this study, the reduction emission was 45% for consistency with Canada's 45% emission reduction goal. As both technologies offer more advantages than the LDAR due to its continuous 24/7 monitoring, data storage and efficient algorithms, it can be concluded that automated detection technologies from the ARPA-E would be more cost-effective than implementing the LDAR with its assumed quarterly check.

Since the upper bound mitigation costs for both IBM and Aeris' technologies were at least higher than the mitigation cost for the LDAR in gas processing plants from Figure 5. However, the mitigation cost for LDAR in compressor stations in Figure 5 were shown to be higher than the calculated mitigation cost in Table 13. If the MONITOR technologies were implemented on the same dataset of compressor stations used by ICF International, the mitigation cost for these technologies would be higher and more cost-effective than the LDAR, considering the advantages offered by the MONITOR technologies.

MONITOR's mitigation costs were then used as an initial guess for each emission source in the optimisation model and the results are tabulated in Table 14 below.

Table 14: Optimised mitigation cost for each emission source

Emission Source	Optimised Mitigation Cost ($CAD/tonne CO2eq)
Clearstone: Reciprocating Compressors	1.809e-03
Clearstone: Separator	2.532e-04
Clearstone: Wellheads	2.918e-03
Clearstone: Dehydrator-Glycol	4.155e-05
Clearstone: Flare Knockout Drum	5.537e-04
Clearstone: Gas Pipeline Header	5.094e-05
Clearstone: Production Tank (Fixed Roof)	1.969e-03

Clearstone: Screw Compressor	8.823e-05
Clearstone: Electric Driver Compressor	5.800e-05
GreenPath: Wellheads	1.094e-02
GreenPath: Separators	6.959e-03

The optimised results indicated that virtually no costs were deducted in mitigating one tonne of CO_2 to reach 45% emission reduction target. However, the calculated mitigation costs far exceeded the optimised results as the mitigation costs were returned as profit for reducing one tonne of CO_2. This was likely that the optimised results were in the local minima instead of the global minima due to the use of the local solver of Brent. The nature of the Brent solver utilised the secant method along with the inverse quadratic interpolation, which resulted in a superlinear convergence. Unfortunately, the nature of the Brent algorithm caused a reduction in the accuracy of the minimisation. Whilst the optimised results seemed sensible due to the very low OPEX per year, it would be more desirable to implement more efficient solvers to reach a global minimum convergence.

Figure 5 showed that majority of the conventional mitigation technologies (except LDAR)'s mitigation costs were between $0-20 Mcf/methane reduced. Which meant that no amount of cost was saved by implementing the conventional technologies. The optimisation results and mitigation cost calculations indicated the savings amount and the efficiency of low-cost methane monitoring technologies in contrast to conventional mitigation technologies including the LDAR. Conclusive that low-cost methane monitoring technologies should be developed more and are subject to further research for companies and academia.

However, the mitigation costs of the MONITOR technologies are subject to uncertainties mainly due to the AECO-C gas price fluctuations. Therefore, it is imperative to discuss the results of the sensitivity analysis and the sensitivity graphs are shown below in Figure 30, 31 and 32.

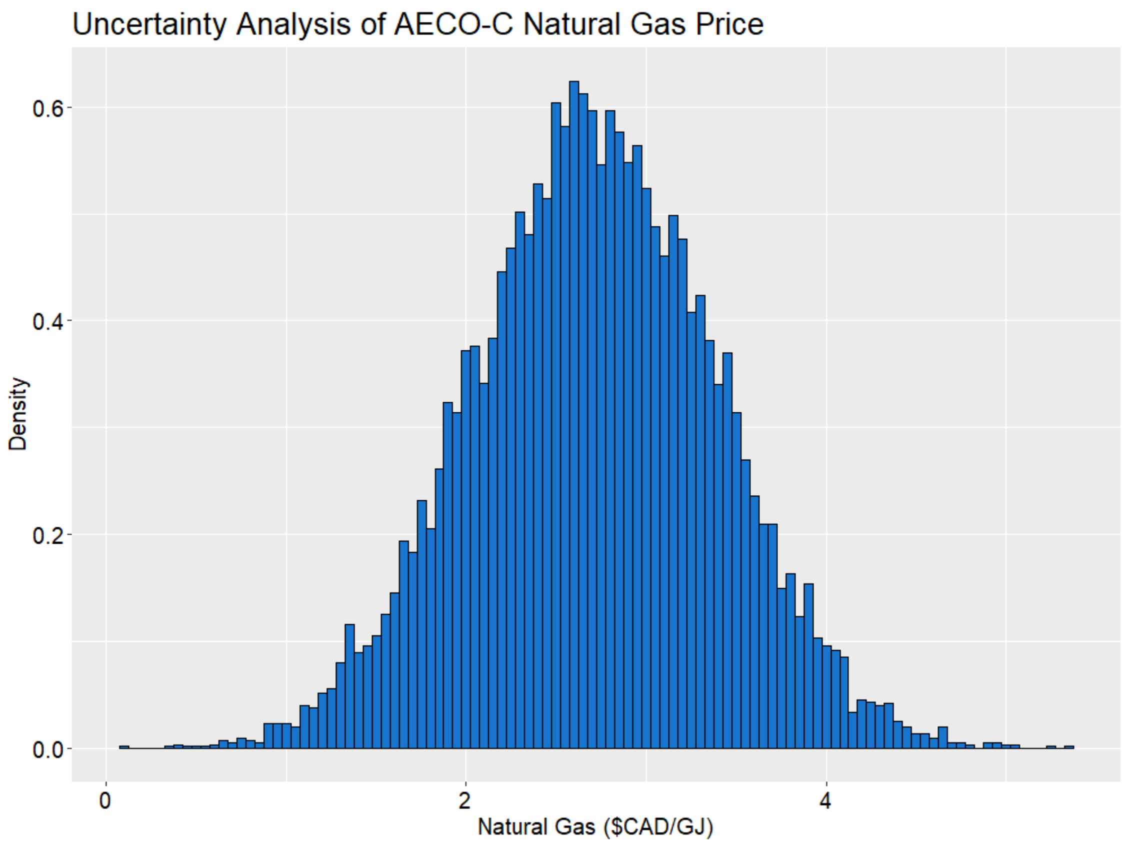

Figure 30: Sensitivity analysis of AECO-C Natural Gas Price

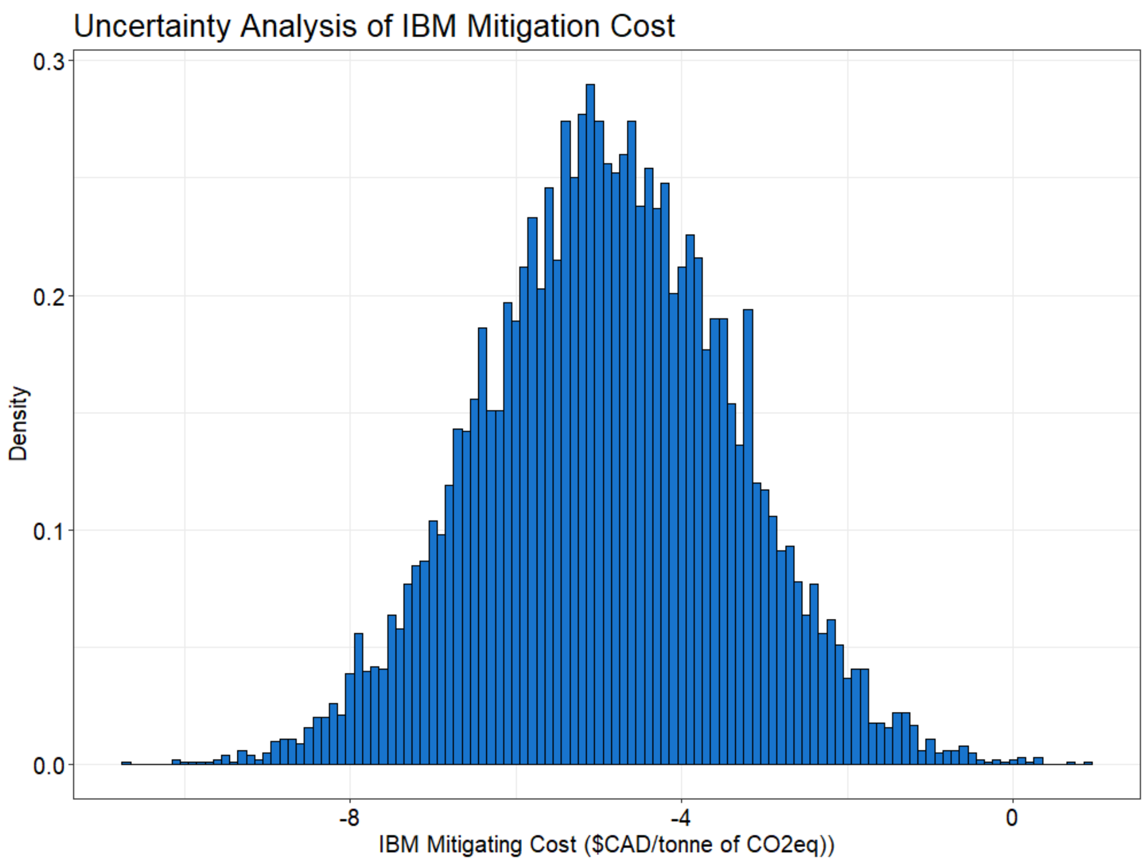

Figure 31: Sensitivity analysis of IBM mitigation cost

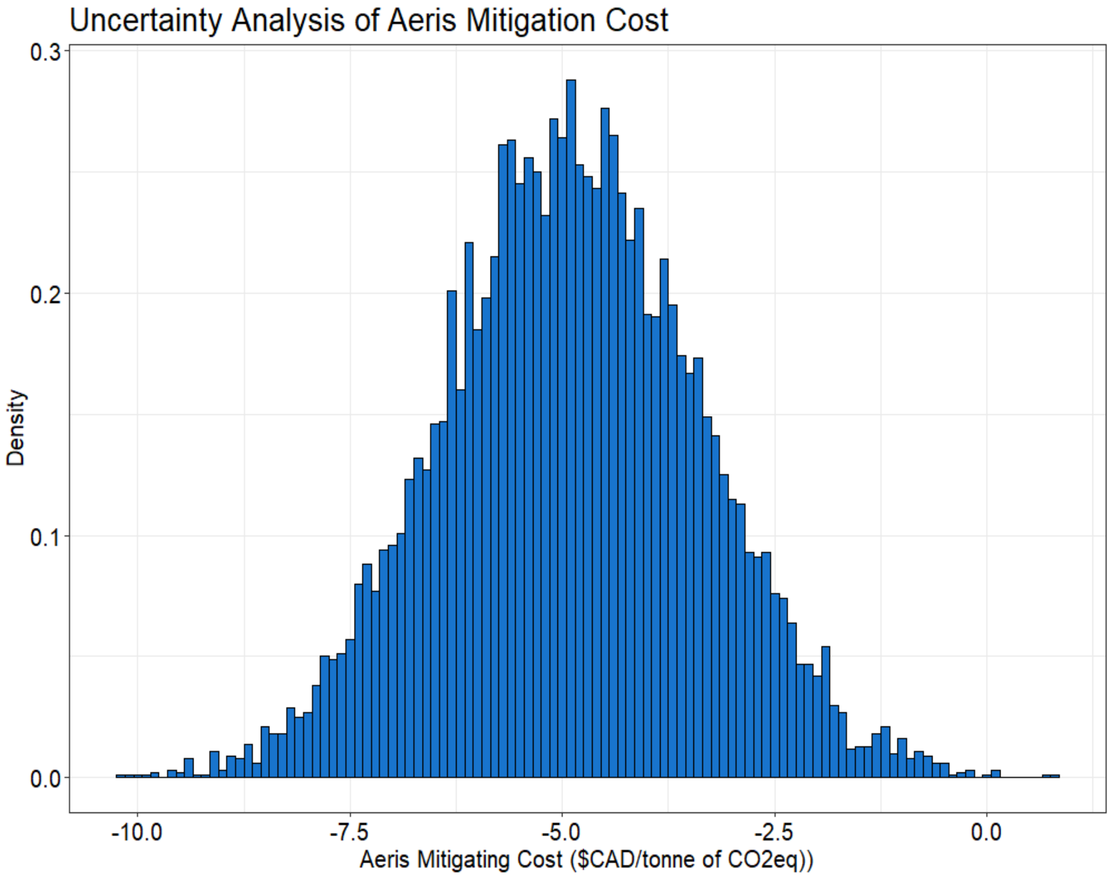

Figure 32: Sensitivity analysis of Aeris mitigation cost

Starting from Figure 30 the AECO-C natural gas price tends to fall between $1.5-4CAD/GJ with approximately a mode of $2.5 CAD/GJ. The AECO-C natural gas price tends to exhibit a value of 2.4-3$CAD/GJ as these prices have the highest density within the sensitivity analysis. The prices were also in line with the base natural gas price from Table 7. It is very unlikely that the natural gas price will fall under $2CAD/GJ or go above $3.5CAD/GJ. However, in 2018, the AECO-C gas price fluctuated immensely from $2.25CAD (2017 price) to $1.48CAD due to scheduled pipeline maintenance, increase in gas production and gas storage changes (Alberta Energy Regulator, 2019). This will be unlikely that the price will decrease again as there are more demands in gas from the mining sectors' oil sands production, which will increase the AECO-C natural gas price. In addition, the natural gas pipeline by TC Energy's maintenance will be finished at the end of 2019 and this influence the increase in AECO-C natural gas price. The projection of the increase in AECO-C gas price is consistent with the high density of the sensitivity analysis in Figure 30.

Figure 31 and 32 shows the sensitivity analysis graphs of the IBM and Aeris' mitigation costs. The graphs considered of the uncertainty of the AECO-C gas price from Figure 30, the mode of the mitigation cost for both technologies were found to be slightly higher than the calculated mitigation cost. The mode mitigation costs were in the range between -$4 and -$6CAD/ tonne

CO_2eq for IBM and the range for Aeris' mitigation cost was -\$3.8 and -\$6.3 CAD/ tonne CO_2eq. It is reasonable to state that the mitigation cost is highly sensitive to the AECO-C natural gas price as both are directly proportional to each other based on Equation 4. The symmetry for all three figures is consistent with each other so it can be inferred the most cost-effective mitigation cost would be heavily reliant on a low AECO-C natural gas price. However, if the AECO-C natural gas price were to increase steadily as in Table 7, it would still be cost-effective as there would still be a guaranteed amount of money to be saved. The AECO-C natural gas price in Table 7 are subject to volatility, hence there is a chance that the prices will expect a substantial increase from the forecasted base price. The high price case was assumed that if the gas production from North America were to be haltered, an increase in the number of LNG projects and strong demand from oil sands production. Should these assumptions to be true, the mitigation cost for MONITOR's technologies will be less cost-effective. Nevertheless, if the volatility in AECO-C natural gas price was to increase or decrease, the mitigation cost for these low-cost, innovative technologies would still prove to be more cost-effective than conventional mitigation technologies due to its low OPEX per year.

Chapter 5: Conclusion

This study details the techno-economic assessment on evaluating the existence of the super-emitters and assessing new, innovative mitigation technologies in hopes of reducing the super-emitters' emissions in a cost-effective way. A range of methane emission sources within the American natural gas supply chain were discussed as a context for the super-emitter study. Prevalent emission sources discussed were wells completion and liquids unloading, as both were founded to have frequent methane emissions across literature. Liquids unloading were a possible super-emitting site due to the high contribution of methane emissions by automatic plunger lift systems, however, these lift systems were sampled differently in contrast to manual lift systems, so recorded emissions were starkly different. This meant that super-emitters are relatively stochastic and can come into existence in any supply chain stage.

Characterisation of methane emissions were discussed including the data collection methodology. Data collection methodology were divided into bottom-up and top-down methods. BU methods are more accurate as the collection of methane emissions were ground-based measurements whilst TD methods were airborne-based. BU methods directly tackle emission sources and can be inferred where the sources of the super-emitters are located. Relevant mitigation technologies can then be implemented appropriately. TD methods overestimates the methane emissions since other methane emissions from different sectors will be aggregated into the measurement and this is one of TD's biggest uncertainties. The definition of super-emitters was discussed, it was concluded that super-emitters are a group of emitters that are characterised by abnormal process conditions and stochastic by nature. The abnormal process conditions are characterised as equipment malfunctions, equipment leaks, inefficient operation of equipment, etc. Conventional mitigation technologies and its operations were extensively discussed and how specific technologies such as the LDAR or low-bleed pneumatic controllers have better reduction opportunities in a more cost-effective way than other mitigation technologies.

The study incorporated a modelling workflow on the techno-economic assessment on the super-emitters and the mitigating options. The best data input for efficient super-emitters analysis was to use a bottom-up dataset with information on the emission source, the reporting segment, emission volumes and type of emissions. However, top-down datasets were also considered due to scarcity of bottom-up datasets and insights could still be derived from its data visualisation.

The technical assessment of this study evaluates the best statistical distribution to characterise methane emissions in order to quantify the heavy-tail characteristic of the distribution curve appropriately. The heavy-tail attribute often hints an indication of super-emitters in the dataset. Weibull, Gamma, log-normal and log-logistic distribution were chosen for their quantification of the heavy-tail characteristic within a dataset. From all the considered datasets, log-logistic distribution was the best distribution. However, the distribution selected in this study for all datasets was log-normal. This is because the nature of the methane emissions from the dataset were not categorical dependent variables, which is an underlying characteristic input for log-logistic distributions. It was appropriate to choose log-normal since the log-distribution curves' qualitative and quantitative results showed similarities amongst each other.

EUR was known to significantly impact the methane emissions. A high methane emission with respect to a low EUR immensely influences the existence of the super-emitters. To model the significance of the EUR against the methane emissions, a Monte Carlo simulation was performed to produce 10,000 cumulative distribution curves. Prior to the simulation, the EUR was best fitted with a log-normal distribution and this was a crucial step to ensure the calculations between the EUR and methane emissions were compatible with each other. The Monte Carlo simulation showed that median emissions from bottom-up datasets were in the range 1.849-33.766% with, 5th percentile of 0.0824-3.576% and 95th percentile of 14.006-100%. Top-down datasets 'median emissions were in the range 15.448-100%, with 5th percentile of 0.569-12.284% and 75th percentile of 39.831-100%. The 95th percentile for the top-down datasets were not considered as the 95th percentile scenarios only shows methane emissions more than 100% of EUR, which these situations were deemed unrealistic. The percentile ranges for both bottom-up and top-down datasets cannot be compared due to their associated uncertainties and differences in emission magnitude. Nevertheless, the variation between the percentiles showed how 10,000 scenarios of EUR variation can influence the existence of the super-emitters greatly, in which proved that the super-emitters problem is prevalent in the natural gas supply chain.

An in-depth analysis of the super-emitters for all the datasets were conducted. This was done by constructing Lorenz curves of the cumulative total population against the cumulative total emissions of a dataset. All the Lorenz curves showed that all the datasets are at least significantly skewed. This was evident as most of the datasets remaining 20% of total population emitted at least 80% of total emissions. This skewness indicated that the contribution was from the super-emitters. However, this type of analysis was unfair as certain facilities have different gas production rates and methane emissions vary as a result. The ideal Lorenz curve analysis is to model the cumulative population of methane emissions as a percentage of instantaneous gas production rate (as "of throughput") instead of cumulative population of an emitter. A good example of this is Figure 23, the difference in skewness from this graph and Figure 22 were stark, especially the skewness contribution. 50% of the remaining total population emitted 82% of total methane emissions, which showed that Figure 22 underestimated the super-emitters' contribution. The main influence was that the instantaneous gas production rate was low whilst the methane emissions were high, a similar reasoning with the EUR. Hence, it is imperative to model methane emissions as "of throughput" basis. To analyse the skewness contribution, the top 5% emitters for all datasets were plotted as scatter graphs against the average absolute emission rate for each respective dataset. Majority of these top 5% emitters' emissions exceeded the average absolute emission rate. From the bottom-up scatter graphs, the top 5% emitters, which can be classified as the super-emitters, were found to be in the range of 5-100m^3/day on average. Top-down's top 5% emitters' range were 105-300m^3/day on average. The visualisation of the top 5% emitters showed that typical super-emitters' emissions were relatively low on average and does not provide the expected significant skewness contribution. However, the significant skewness contribution comes from emissions outside the average range, where the magnitude for these emissions were extremely large. These extreme emissions were in the range of 100-1600m^3/day for bottom-up datasets and 300-2500m^3/day for top-down datasets. It was clear that significant skewness contribution did not come from the low-average emissions but from emissions that immensely exceeded the average emissions.

Mitigation technologies provided by IBM and Aeris Technologies, under ARPA-E's MONITOR program, were economically assessed in a 45% emission reduction scenario. Mitigation costs for both technologies were at -$2.187CAD/tonne CO_2eq and -$3.149 CAD/ tonne CO_2eq. As both mitigation costs are negative, this indicates the amount of money being saved rather than spent. Whilst the mitigation costs calculations were site specific, it showed how cost-effective the MONITOR technologies are since the required OPEX are extremely cheap in contrast to conventional mitigation technologies. An optimisation model was executed to evaluate the optimised mitigation cost. The results from the optimisation model showed that the mitigation costs are virtually free, however, the calculated mitigation costs showed better results than the optimised results. The reasons for this was the optimised results were obtained in local minima due to the limited choice and ability of the algorithm solvers in RStudio. Nevertheless, the optimised results still proved to be relatively sensible as the considered mitigation technologies were extremely efficient enough to reduce the methane emissions.

Sensitivity analysis for the AECO-C natural gas price and the mitigation costs were conducted to model cost fluctuations in 10 years. Results from the sensitivity analysis showed that mitigation costs increase as the AECO-C natural gas price was expected to increase as well, therefore mitigation of methane emissions and the super-emitters would be more profitable. However, should the AECO-C natural gas price decrease, the mitigation cost would decrease but it would still be cost-effective as the results from the sensitivity analysis showed that the costs are negative regardless of the AECO-C natural gas price's volatility.

Possibility of future work from this study is to replace the proxy EUR data with EUR data from Canada to create more accurate results from the Monte Carlo simulations. It is more desirable to model the Monte Carlo simulation with EUR from specific well types such as gas, shale and coalbed methane. It is likely that the EUR varies based on well types and it would be a valuable study to conduct Monte Carlo simulations based on EUR with different well types. In addition, more methane emissions data should be collected and comply with the natural gas supply chain in its respective region. Skewness contributions can vary based on methane emissions and instantaneous gas production rate. Companies and academia undertaking a super-emitter study must record the instantaneous gas production rate and plot the Lorenz curves on a "of throughput" basis to prevent making biased conclusions for certain emitting sites. Not only that, the most valuable dataset to identify super-emitters effectively is a bottom-up type dataset. This is because methane emissions are collected as a ground-based measurement, so the emission source, type of emission and reporting segments can be recorded accurately in contrast to top-down type datasets.

Another future work is to conduct a more detailed economic assessment and assess more MONITOR type technologies to reduce emissions. It was evident that these low-cost, innovative technologies proved to be more cost-effective than conventional technologies and should be explored further.

References

Aitkin, M. *et al.* (2009) *Statistical Modelling in R*. New York, United States: Oxford University Press Inc.

Alberta Energy Regulator (2019) *AECO-C Price*. Available at: https://www.aer.ca/providing-information/data-and-reports/statistical-reports/st98/prices-and-capital-expenditure/natural-gas-prices/aeco-c-price (Accessed: 14 September 2019).

Allen, D. T. *et al.* (2013) *Measurements of Methane Emissions at Natural Gas Production Sites in the United States*. doi: 10.1073/pnas.1304880110.

Allen, David T *et al.* (2015) 'Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Liquid Unloadings', *Environmental Science & Technology*, 49, pp. 641–648. Available at: http://dept.ceer.utexas.edu/methane2/study/docs/UT Study Unloadings FAQ to SC.pdf.

Allen, David T. *et al.* (2015) 'Methane emissions from process equipment at natural gas production sites in the United States: Pneumatic controllers', *Environmental Science and Technology*, 49(1), pp. 633–640. doi: 10.1021/es5040156.

Alvarez, R. A. *et al.* (2018) 'Assessment of methane emissions from the U.S. oil and gas supply chain', *Science*, 188(July), pp. 186–188. doi: 10.1126/science.aar7204.

Anderson, T. W. and Darling, D. A. (1952) 'Asymptotic theory of certain "Goodness of Fit" criteria based on stochastic processes', *The Annals of Mathematical Statistics*, 23, pp. 193–212.

ARPA-E (2015) *Methane Leak Detection System*. Available at: https://arpa-e.energy.gov/?q=slick-sheet-project/methane-leak-detection-system (Accessed: 14 September 2019).

Balcombe, P. *et al.* (2015) *Methane and CO2 Emissions From the Natural Gas Supply Chain An Evidence Assessment*. London.

Balcombe, P. *et al.* (2017) 'The Natural Gas Supply Chain: The Importance of Methane and Carbon Dioxide Emissions', *ACS Sustainable Chemistry and Engineering*, 5(1), pp. 3–20. doi: 10.1021/acssuschemeng.6b00144.

Balcombe, P. *et al.* (2018) 'Methane emissions: choosing the right climate metric and time horizon', *Environmental Science: Processes and Impacts*, 20(10), pp. 1323–1339. doi: 10.1039/c8em00414e.

Balcombe, P., Brandon, N. P. and Hawkes, A. D. (2018) 'Characterising the distribution of methane and carbon dioxide emissions from the natural gas supply chain', *Journal of Cleaner Production*. Elsevier Ltd, 172(2018), pp. 2019–2032. doi: 10.1016/j.jclepro.2017.11.223.

Balcombe, P., Brandon, N. P. and Hawkes, A. D. (2019) 'Characterising the distribution of methane and carbon dioxide emissions from the natural gas supply chain', *Journal of Cleaner Production*. Elsevier Ltd, 172(2018), pp. 2019–2032. doi: 10.1016/j.jclepro.2017.11.223.

Brandt, A. *et al.* (2014) 'Methane Leaks from North American Natural Gas Systems', *Science*, 343(6172), pp. 733–735. doi: 10.1126/science.1247045.

Brandt, A. R., Heath, G. A. and Cooley, D. (2016) 'Methane Leaks from Natural Gas Systems Follow Extreme Distributions', *Environmental Science and Technology*, 50(22), pp. 12512–12520. doi: 10.1021/acs.est.6b04303.

CCAC (2016) *Reducing Methane Emissions*.

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

Delignette-Muller, M.-L. *et al.* (2019) *Package ' fitdistrplus '*.

EPA (2006a) *Cost-effective directed inspection and maintenance control opportunities at five gas processing plants and upstream gathering compressor stations and well sites*.

EPA (2006b) *Lessons Learned Installing Vapor Recovery Units on Storage Tanks Installing Vapor Recovery Units on Storage Tanks*. Available at: https://www.epa.gov/sites/production/files/2016-06/documents/ll_final_vap.pdf.

EPA (2006c) *Lessons Learned Options For Reducing Methane Emissions From Pneumatic Devices In The Natural Gas Industry*.

EPA (2006d) *Lessons Learned Replacing Gas-Assisted Glycol Pumps with Electric Pumps Replace Gas-Assisted Glycol Pumps With Electric Pumps*. Available at: https://www.epa.gov/sites/production/files/2016-06/documents/ll_glycol_pumps3.pdf.

EPA (2009) *Lessons Learned from Natural Gas STAR Partners Reduced Emissions Completions for Hydraulically Fractured Natural Gas Wells*. Available at: https://www.epa.gov/sites/production/files/2016-06/documents/reduced_emissions_completions.pdf.

EPA (2014) *Wet Seal Degassing Recovery System for Centrifugal Compressors*. Available at: https://www.epa.gov/sites/production/files/2016-06/documents/capturemethanefromcentrifugalcompressionsealoildegassing.pdf.

Goulet, V. *et al.* (2019) *Package ' actuar '*.

Government of Alberta (2019) *Reducing Methane Emissions*. Available at: https://www.alberta.ca/climate-methane-emissions.aspx (Accessed: 11 August 2019).

GreenPath Energy Ltd. (2016) *GreenPath 2016 Alberta Fugitive and Vented Emissions Inventory*.

GreenPath Energy Ltd. (2017) *HISTORICAL CANADIAN FUGITIVE EMISSIONS MANAGEMENT*.

Guo, K. *et al.* (2017) 'Characteristic Production Decline Patterns for Shale Gas Wells in Barnett', *International Journal Sustainable Future for Human Security*, 5(1), pp. 12–21.

IBM (2019) 'An Intelligent Methane Management Solution (AIMS)', p. 18.

ICF International (2014) *Economic Analysis of Methane Emission Reduction Opportunities in the U.S. Onshore Oil and Natural Gas Industries*. doi: 10.1016/S0065-2113(08)60504-0.

de Leeuw, J. and Akaike, H. (1973) *Introduction to Akaike (1973): Information theory as an extension of the maximum likelihood principle*.

Mandler, B. and Allison, E. (2018) *Methane Emissions in the Oil and Gas Industry*, *Petroleum and the Environment*.

Minitab (2019) *The Anderson-Darling statistic*. Available at: https://support.minitab.com/en-us/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/normality/the-anderson-darling-statistic/ (Accessed: 4 August 2019).

O'Connell, E. *et al.* (2019) 'Methane emissions from contrasting production regions within Alberta , Canada : Implications under incoming federal methane regulations', (2018).

OAQPS (2014) *Oil and Natural Gas Sector Liquids Unloading Processes*.

OFX (2019) *Yearly Average Rates*. Available at: https://www.ofx.com/en-gb/forex-news/historical-exchange-rates/yearly-average-rates/ (Accessed: 14 September 2019).

Rubinstein, R. and Kroese, D. (2011) *Simulation and the Monte Carlo Method*. Second. Hoboken, New Jersey: John Wiley & Sons, Inc. doi: 10.1002/9780470230381.

Schwarz, G. (1978) 'Estimating the dimension of a model', *The Annual of Statistics*, 6, pp. 461–464.

Skone, T. J., Littlefield, J. and Marriott, J. (2011) *Life cycle greenhouse gas inventory of natural gas extraction, delivery and electricity production.*, *National Energy Technology Laboratory*. doi: 10.1021/es501997p.

Snipes, M. and Taylor, D. C. (2014) 'Model selection and Akaike Information Criteria : An example from wine ratings and prices', *Wine Economics and Policy*. Elsevier, 3(1), pp. 3–9. doi: 10.1016/j.wep.2014.03.001.

Stamford, L. and Azapagic, A. (2014) 'Life cycle environmental impacts of UK shale gas', *Applied Energy*. Elsevier Ltd, 134, pp. 506–518. doi: 10.1016/j.apenergy.2014.08.063.

Umeozor, E. *et al.* (2019) *Economic and Environmental Impacts of Methane Emissions Reduction In The Natural Gas Supply Chain*. Canada.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Third. Springer.

Wickham, H. (2019) *Package ' dplyr '*. Houston,Texas.

Zavala-araiza, D. *et al.* (2015) 'Allocating Methane Emissions to Natural Gas and Oil Production from Shale Formations', *ACS Sustainable Chemical Engineering*. doi: 10.1021/sc500730x.

Zavala-Araiza, D., Lyon, D. R., *et al.* (2015) 'Reconciling divergent estimates of oil and gas methane emissions', *PNAS*, 112(51), pp. 15597–15602. doi: 10.1073/pnas.1522126112.

Zavala-Araiza, D., Lyon, D., *et al.* (2015) 'Toward a Functional Definition of Methane Super-Emitters: Application to Natural Gas Production Sites', *Environmental Science & Technology*, 49(13), pp. 8167–8174. doi: 10.1021/acs.est.5b00133.

Zavala-Araiza, D. *et al.* (2017) 'Super-emitters in natural gas infrastructure are caused by abnormal process conditions', *Nature Communications*, 8, pp. 1–10. doi: 10.1038/ncomms14012.

Zavala-Araiza, D. *et al.* (2018) *Methane emissions from oil and gas production sites in Alberta, Canada*, *Elementa Science of the Anthropocene*. Available at: https://www.elementascience.org/article/10.1525/elementa.284/.

Appendices (Literature Review)

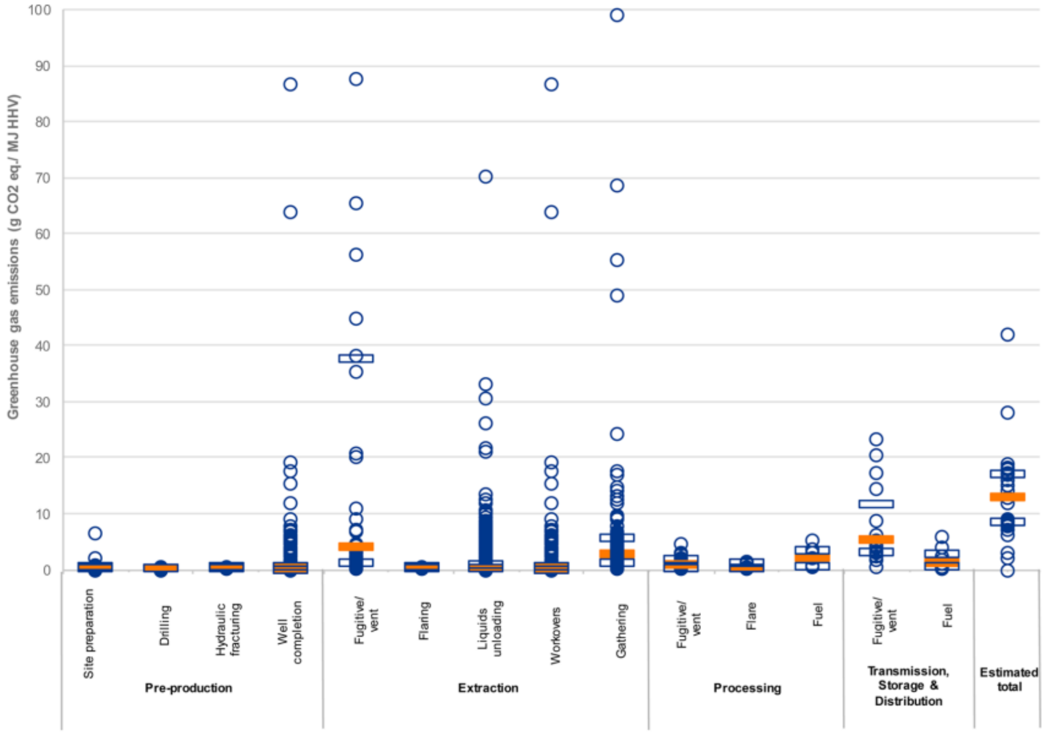

Figure 33: Estimated supply chain greenhouse gas emissions with the range from 0 – 100 gCO₂eq/HJ MMV. A circle is defined as the literature estimate. The orange bar indicated the median estimate of a stage in the supply chain. Hollow bars (shown for some stages) Source: (Balcombe *et al.*, 2017)

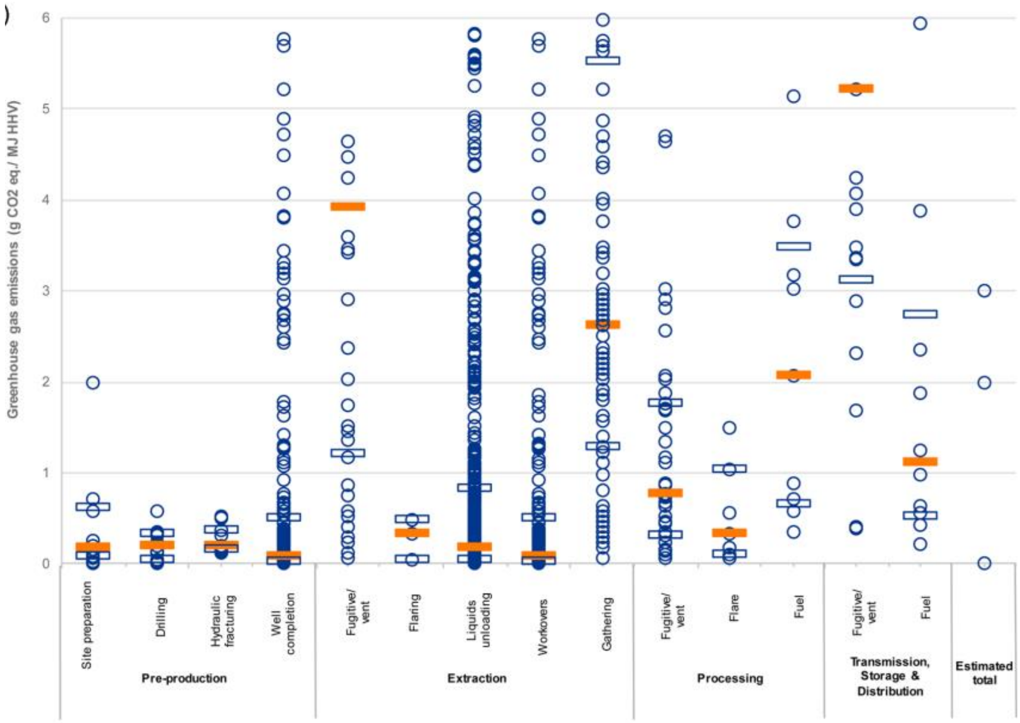

Figure 34: Estimated supply chain greenhouse gas emissions with the range from 0 – 6 gCO_2eq/ MJ HHV. A circle is defined as the literature estimate. The orange bar indicated the median estimate of a stage in the supply chain. Hollow bars (shown for some stages) Source: (Balcombe *et al.*, 2017)

well type	equipment	data	sample size	emissions (1000 m^3 CH_4/ completion)			
				mean	median	min	max
conventional		primary	10	4.9	5.7	0.0	7.4
		secondary	8	0.9	1.0	0.0	2.0
unconventional	RECs	primary	76	3.0	1.1	0.0	24.9
		secondary	14	39.3	15.0	0.0	210.1
	non-RECs	primary	88	11.9	5.8	0.3	100.1
		secondary	73	606.0	245.8	1.3	6800.0

Figure 35: Methane emissions data on wells completion. Primary data sources are measured methane emissions whilst secondary data sources are estimated or based on other literature. Source: (Balcombe *et al.*, 2017)

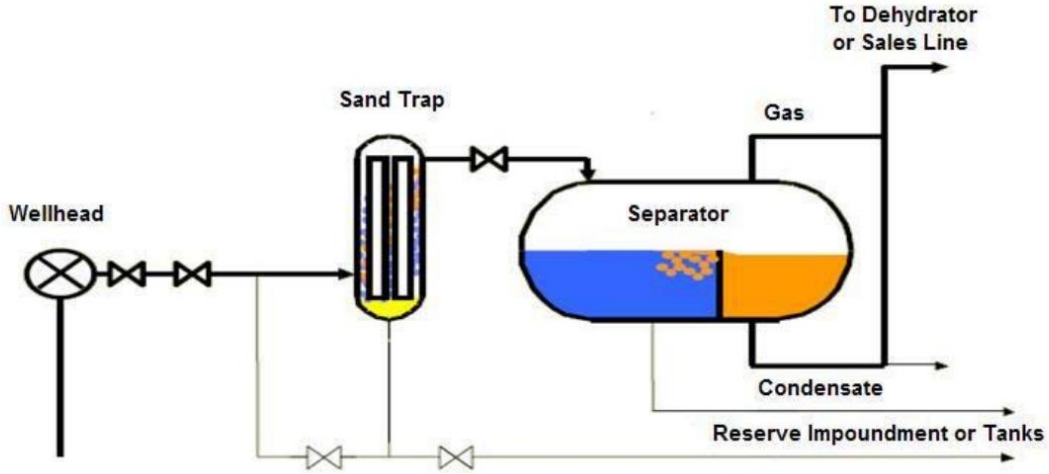

Figure 36: Schematic of REC equipment. Source: (EPA, 2009)

Table 15: Cost benefit analysis for mitigation technologies. Source: (ICF International, 2014)

Technologies	CAPEX	OPEX	Revenue	Payback time
Plunger lifts	$2400 - $9600	$860-$1600	$18800 - $73000	1 to 7 months
Dry seals	$400,000	$17,500	$180,500	29 months
Wet seal degassing recovery systems	$33,700	Minimal	$120,000	3 months
Vapour recovery units	$44500 - $129000	$9100 - $20800	$20,000 - $380000	4 to 51 months
Low-bleed pneumatic conversion	$260 to $420 (End of life approach)	N/A	$200 to $800 (End of life approach)	4 to 25 months (End of life approach)
	$2300(Early replacement)		$1040(Early replacement)	27 months (Early replacement)
	$840 (Retrofit)		$920 (Retrofit)	11 months (Retrofit)
Rod packing replacement	$2000 every 3 years	N/A	$3500	7 months
	$670 per year			
Electric pumps	$1200 - $12000	$120 - $1200 (Electricity)	$10000	1 to 16 months
LDAR	$124,000 (IR Camera)	$10 - $6900 (Repair costs)	$1500	Dependent on repair costs

Appendices (Methodology/Results)

Table 16: Clearstone dataset input

Reporting Segment	Emission Type	Emission Source	Emission Volume (m^3/day)
Gas Multiwell effluent	Leak	Reciprocating Compressor	0.762
Gas Multiwell effluent	Leak	Reciprocating Compressor	4.230
Gas Multiwell effluent	Leak	Reciprocating Compressor	12.860
Gas Multiwell effluent	Leak	Reciprocating Compressor	13.845
Gas Multiwell Group	Leak	Separator	3.188
Gas Multiwell Group	Leak	Separator	3.791
Gas Multiwell Group	Leak	Separator	0.719
Gas Multiwell Group	Leak	Separator	1.925
Gas Multiwell proration outside SE AB	Leak	Separator	0.383
Gas Multiwell proration outside SE AB	Leak	Separator	0.383
Gas Multiwell proration outside SE AB	Leak	Reciprocating Compressor	1.915
Gas Multiwell proration outside SE AB	Leak	Reciprocating Compressor	15.321
Gas Multiwell proration outside SE AB	Leak	Reciprocating Compressor	0.766
Gas Multiwell proration outside SE AB	Leak	Dehydrator - Glycol	1.915
Gas Multiwell proration outside SE AB	Vent	Production Tank (fixed roof)	5.209
Gas Multiwell proration outside SE AB	Leak	Reciprocating Compressor	1.149
Gas Multiwell proration outside SE AB	Leak	Reciprocating Compressor	0.383
Gas Multiwell proration outside SE AB	Leak	Reciprocating Compressor	1.915
Gas Multiwell proration outside SE AB	Leak	Reciprocating Compressor	1.149
Gas Multiwell effluent	Leak	Screw Compressor - Electric Driver	0.364
Gas Multiwell proration outside SE AB	Leak	Separator	5.516
Gas Multiwell proration outside SE AB	Vent	Production Tank (fixed roof)	34.725
Gas Multiwell proration outside SE AB	Leak	Flare KnockOut Drum	1.460
Gas Multiwell proration outside SE AB	Vent	Production Tank (fixed roof)	18.071

Gas Multiwell effluent	Leak	Separator	1.089
Gas Multiwell effluent	Leak	Wellhead	0.726
Gas Multiwell effluent	Leak	Wellhead	5.443
Gas Multiwell effluent	Leak	Wellhead	0.726
Gas Multiwell effluent	Leak	Wellhead	1.814
Gas Multiwell effluent	Leak	Wellhead	2.903
Gas Multiwell effluent	Leak	Wellhead	10.887
Gas Multiwell effluent	Leak	Wellhead	1.089
Compressor Station	Leak	Reciprocating Compressor	13.433
Compressor Station	Leak	Reciprocating Compressor	57.572
Compressor Station	Leak	Reciprocating Compressor	1.979
Compressor Station	Leak	Separator	2.085
Compressor Station	Leak	Separator	4.075
Compressor Station	Leak	Separator	0.695
Compressor Station	Leak	Separator	6.951
Compressor Station	Vent	Production Tank (fixed roof)	4.982
Compressor Station	Leak	Reciprocating Compressor	3.855
Compressor Station	Leak	Screw Compressor	4.543
Compressor Station	Leak	Screw Compressor	3.029
Gas Gathering System	Leak	Reciprocating Compressor	25.982
Gas Gathering System	Leak	Reciprocating Compressor	8.547
Gas Gathering System	Leak	Reciprocating Compressor	80.900
Compressor Station	Leak	Reciprocating Compressor	5.367
Compressor Station	Leak	Other (describe in notes)	1.440
Gas Gathering System	Leak	Reciprocating Compressor	1.740
Gas Gathering System	Leak	Reciprocating Compressor	0.348
Gas Gathering System	Leak	Reciprocating Compressor	3.079
Gas Gathering System	Leak	Dehydrator - Glycol	1.368
Gas Gathering System	Leak	Dehydrator - Glycol	2.052
Gas Gathering System	Leak	Dehydrator - Glycol	4.447
Gas Gathering System	Vent	Production Tank (fixed roof)	3.357
Gas Gathering System	Leak	Reciprocating Compressor	36.509
Gas Gathering System	Unlit Flare	Flare KnockOut Drum	84.214
Gas Gathering System	Leak	Reciprocating Compressor	2.445
Gas Gathering System	Leak	Screw Compressor	11.558
Gas Gathering System	Leak	Dehydrator - Glycol	1.010
Gas Gathering System	Unlit Flare	Flare KnockOut Drum	32.797
Gas Gathering System	Leak	Reciprocating Compressor - Electric Driver	15.446
Gas Gathering System	Leak	Reciprocating Compressor - Electric Driver	6.865
Gas Gathering System	Leak	Pig Trap	7.285
Gas Gathering System	Leak	Gas Pipeline Header	3.122
Gas Gathering System	Leak	Reciprocating Compressor	12.142

Gas Gathering System	Leak	Reciprocating Compressor	3.434
Gas Gathering System	Leak	Production Tank (fixed roof)	165.726
Gas Gathering System	Leak	Reciprocating Compressor	0.707
Gas Gathering System	Leak	Reciprocating Compressor	0.353
Gas Gathering System	Leak	Reciprocating Compressor	0.353
Gas Gathering System	Leak	Reciprocating Compressor	12.014
Gas Gathering System	Leak	Gas Pipeline Header	1.750
Gas Gathering System	Leak	Reciprocating Compressor	22.553
Gas Gathering System	Leak	Reciprocating Compressor	1.428
Gas Gathering System	Leak	Gas Pipeline Header	1.037
Gas Gathering System	Leak	Gas Pipeline Header	0.346
Gas Gathering System	Leak	Gas Pipeline Header	0.692
Gas Gathering System	Leak	Reciprocating Compressor	14.263
Gas Gathering System	Leak	Reciprocating Compressor	5.851
Gas Gathering System	Leak	Reciprocating Compressor	1.413
Gas Gathering System	Leak	Reciprocating Compressor	3.288
Gas Gathering System	Leak	Reciprocating Compressor	1.097
Gas Gathering System	Leak	Reciprocating Compressor	11.805
Gas Gathering System	Leak	Reciprocating Compressor	15.315
Gas Gathering System	Leak	Gas Pipeline Header	4.124
Gas Gathering System	Leak	Dehydrator - Glycol	0.692
Gas Gathering System	Leak	Dehydrator - Glycol	2.069
Gas Gathering System	Leak	Flare KnockOut Drum	0.333
Gas Gathering System	Leak	Separator	2.099
Gas Gathering System	Leak	Screw Compressor	0.335
Gas Gathering System	Leak	Screw Compressor	0.335
Gas Gathering System	Leak	Screw Compressor	0.663
Gas Gathering System	Leak	Screw Compressor	0.663
Gas Gathering System	Leak	Reciprocating Compressor	0.356
Gas Gathering System	Leak	Reciprocating Compressor	16.016
Gas Gathering System	Leak	Dehydrator - Glycol	0.354
Gas Gathering System	Leak	Reciprocating Compressor	0.745
Gas Gathering System	Leak	Reciprocating Compressor	1.738
Gas Gathering System	Leak	Flare KnockOut Drum	1.044
Gas Gathering System	Leak	Flare KnockOut Drum	1.044
Compressor Station	Vent	Reciprocating Compressor	24.515
Compressor Station	Leak	Reciprocating Compressor	4.798
Gas Gathering System	Leak	Reciprocating Compressor	19.844
Gas Gathering System	Leak	Reciprocating Compressor	57.262
Gas Gathering System	Unlit Flare	Flare KnockOut Drum	49.819
Gas Gathering System	Leak	Reciprocating Compressor	8.568
Gas Gathering System	Leak	Other (describe in notes)	0.333
Gas Gathering System	Leak	Reciprocating Compressor - Electric Driver	1.659

Gas Gathering System	Leak	Reciprocating Compressor - Electric Driver	1.659
Gas Gathering System	Leak	Reciprocating Compressor - Electric Driver	0.664
Compressor Station	Vent	Production Tank (fixed roof)	360.216
Compressor Station	Leak	Screw Compressor	8.620
Compressor Station	Leak	Separator	1.086
Compressor Station	Leak	Reciprocating Compressor	5.796
Gas Single	Leak	Wellhead	0.342
Gas Multiwell Group	Leak	Wellhead	10.619
Gas Multiwell Group	Leak	Centrifugal Compressor	5.779
Gas Multiwell Group	Vent	Wellhead	0.679
Gas Multiwell Group	Leak	Wellhead	0.336
Gas Multiwell Group	Leak	Separator	3.600
Gas Multiwell proration outside SE AB	Leak	Gas Pipeline Header	6.885
Gas Multiwell Group	Leak	Wellhead	63.272
Gas Multiwell Group	Leak	Separator	1.032
Gas Multiwell Group	Leak	Separator	0.344
Gas Multiwell Group	Leak	Separator	0.688
Gas Multiwell Group	Leak	Separator	0.347
Gas Multiwell Group	Leak	Separator	0.351
Gas Single	Leak	Production Tank (fixed roof)	7.310
Gas Multiwell effluent	Leak	Wellhead	0.349
Gas Multiwell effluent	Leak	Wellhead	1.048
Gas Multiwell Group	Vent	Production Tank (fixed roof)	8.352
Gas Multiwell Group	Leak	Wellhead	75.828
Gas Multiwell proration outside SE AB	Leak	Wellhead	8.166
Gas Multiwell Group	Leak	Wellhead	2.794
Gas Multiwell effluent	Leak	Wellhead	1.414
Gas Multiwell Group	Leak	Wellhead	3.795
Gas Multiwell Group	Leak	Wellhead	0.340
Gas Multiwell Group	Leak	Separator	0.681
Gas Multiwell Group	Leak	Wellhead	1.691
Gas Multiwell Group	Vent	Production Tank (fixed roof)	8.077
Gas Multiwell Group	Leak	Separator	7.990
Gas Multiwell Group	Vent	Separator	8.042
Gas Multiwell Group	Leak	Wellhead	1.399
Gas Multiwell Group	Leak	Wellhead	1.049
Gas Multiwell Group	Leak	Wellhead	41.933
Gas Multiwell Group	Leak	Wellhead	3.946
Gas Multiwell Group	Leak	Wellhead	3.129
Gas Single	Leak	Wellhead	8.343

Gas Multiwell effluent	Leak	Wellhead	1.079
Gas Multiwell Group	Leak	Wellhead	6.593
Gas Multiwell Group	Leak	Separator	0.630
Gas Multiwell proration outside SE AB	Leak	Other (describe in notes)	1.858
Gas Multiwell proration outside SE AB	Leak	Other (describe in notes)	1.485
Gas Multiwell proration outside SE AB	Leak	Other (describe in notes)	3.712
Gas Multiwell proration outside SE AB	Leak	Other (describe in notes)	1.485
Gas Multiwell proration outside SE AB	Leak	Separator	1.114
Gas Multiwell Group	Leak	Wellhead	1126.646
Gas Multiwell Group	Leak	Well Pump	1.408
Gas Multiwell Group	Leak	Wellhead	0.704
Gas Multiwell Group	Leak	Separator	0.704
Gas Multiwell Group	Leak	Separator	1.042
Gas Multiwell effluent	Leak	Wellhead	3.855
Gas Multiwell effluent	Leak	Wellhead	3.833
Gas Multiwell Group	Leak	Wellhead	1.793
Gas Multiwell Group	Leak	Liquid Pump	0.350
Gas Multiwell Group	Leak	Wellhead	0.349
Gas Multiwell Group	Leak	Wellhead	0.349
Gas Multiwell Group	Leak	Separator	1.048
Gas Multiwell effluent	Leak	Wellhead	1.060
Gas Single	Leak	Separator	6.277
Gas Single	Leak	Separator	10.462
Gas Single	Leak	Separator	0.349
Gas Single	Leak	Separator	3.487
Gas Single	Unlit Flare	Flare KnockOut Drum	4.911
Gas Single	Leak	Separator	3.890
Gas Multiwell Group	Leak	Wellhead	1.042
Gas Multiwell Group	Leak	Separator	1.389
Gas Single	Leak	Reciprocating Compressor	2.540
Gas Single	Leak	Reciprocating Compressor	52.980
Gas Single	Vent	Production Tank (fixed roof)	1.647
Gas Single	Leak	Wellhead	0.690
Gas Single	Leak	Wellhead	5.724
Gas Multiwell Group	Leak	Separator	4.719
Gas Multiwell Group	Vent	Wellhead	27.301
Gas Multiwell Group	Leak	Separator	6.878
Gas Single	Vent	Production Tank (fixed roof)	0.737
Gas Multiwell effluent	Leak	Line Heater	0.699

Gas Multiwell effluent	Leak	Line Heater	1.397
Gas Multiwell effluent	Leak	Wellhead	0.699
Gas Single	Leak	Production Tank (fixed roof)	0.000
Gas Multiwell Group	Leak	Wellhead	0.340
Gas Multiwell effluent	Leak	Wellhead	0.363
Gas Multiwell proration outside SE AB	Leak	Wellhead	10.417
Gas Multiwell effluent	Leak	Wellhead	1.076

Table 17: GreenPath dataset input

Reporting Segment	Emission Type	Emission Source	Emission Volume (m^3/day)
Gas Well	Leak	Filter/Separation	7.628
Gas Single-well Battery	Leak	Tankage	1525.608
Gas Well	Leak	Filter/Separation	13.222
Gas Well	Leak	Wellhead	2.543
Gas Well	Leak	Wellhead	2.543
Gas Well	Leak	Filter/Separation	1.526
Gas Well	Leak	Wellhead	20.341
Gas Multiwell Proration Outside Se Alberta Battery	Leak	Filter/Separation	2.034
Gas Multiwell Proration Outside Se Alberta Battery	Leak	Compression	1.526
Gas Multiwell Proration Outside Se Alberta Battery	Leak	Wellhead	2.034
Gas Well	Leak	Wellhead	6.611
Gas Well	Leak	Wellhead	7.628
Gas Well	Leak	Wellhead	6.102
Gas Multiwell Group Battery	Leak	Filter/Separation	8.137
Gas Multiwell Proration Outside Se Alberta Battery	Leak	Wellhead	1.017
Gas Multiwell Proration Outside Se Alberta Battery	Leak	Wellhead	3.051
Gas Multiwell Proration Outside Se Alberta Battery	Leak	Wellhead	0.509
Gas Multiwell Proration Outside Se Alberta Battery	Leak	Wellhead	1.017

Gas Multiwell Proration Outside Se Alberta			
Battery	Vent	Compression	2.543
Gas Well	Leak	Wellhead	2.543
Gas Well	Leak	Filter/Separation	25.427
Gas Well	Leak	Filter/Separation	2.543
Gas Well	Leak	Filter/Separation	2.543
Gas Well	Leak	Filter/Separation	5.085
Compressor Station	Vent	Compression	25.427
Gas Well	Leak	Filter/Separation	7.628

Table 18: Zavala-Araiza et al dataset input

Reporting Segment	Gas Production Rate (m^3/day)	Emission Volume (m^3/day)
Production	1105.36	1.5
Production	713.363	22.5
Production	935.8475	8.7
Production	1493.825	0.6
Production	194.2325	5.4
Production	759.2725	19.8
Production	709.8315	6.6
Production	1779.876	3.9
Production	678.048	6.9
Production	42.378	3.6
Production	4040.036	18.9
Production	1663.337	4.2
Production	148.323	2399.7
Production	12709.87	16.5
Production	5138.333	113.1
Production	1938.794	16.2
Production	4816.966	99.6
Production	0	1133.7
Production	1878.758	168
Production	0	6.3
Production	14.126	1.8
Production	7695.139	54.9
Production	7695.139	52.5
Production	0	8.7
Production	17.6575	7.5
Production	5848.164	35.1
Production	7603.32	288
Production	1048.856	23.7
Production	1048.856	18.3
Production	6060.054	409.8
Production	6060.054	103.8

Production	4760.462	518.1
Production	5706.904	310.5
Production	14253.13	432.6
Production	10894.68	115.8
Production	2948.803	152.4
Production	4594.482	93.9
Production	7267.827	75
Production	0	660.9
Production	1762.219	17.4
Production	187.1695	410.7
Production	2115.369	75.9
Production	3386.709	29.1
Production	1126.549	91.8
Production	1126.549	26.1
Production	4287.241	96.9
Production	8016.505	111.9
Production	8016.505	44.4
Production	18459.15	25.2
Production	374.339	40.2
Production	3266.638	8.7
Production	0	12.3
Production	685.111	124.8
Production	18088.34	639
Production	23131.33	24.9
Production	0	6.3
Production	16244.9	258
Production	3227.791	47.1
Production	2839.326	189.9
Production	2839.326	164.4

Table 19: O'Connell et al dataset input

Reporting Segment	Minimum Emission Volume (m^3/day)	Maximum Emission Volume (m^3/day)	Mean Emission Volume (m^3/day)	Median Emission Volume (m^3/day)
Production	50.848	1219.113	820.170	666.608
Production	73.962	507.844	290.903	290.903
Production	228.885	268.639	248.762	248.762
Production	21.972	457.683	199.226	124.976
Production	72.004	150.289	107.891	101.381
Production	96.453	96.453	96.453	96.453
Production	11.747	155.705	83.726	83.726
Production	20.815	141.968	81.392	81.392
Production	41.387	123.624	78.667	60.893
Production	18.474	131.630	75.052	75.052

Production	30.509	117.874	74.191	74.191
Production	10.937	136.660	73.798	73.798
Production	67.834	67.834	67.834	67.834
Production	35.680	71.228	55.962	58.469
Production	55.945	55.945	55.945	55.945
Production	43.315	61.030	52.173	52.173
Production	10.470	98.464	51.060	44.245
Production	43.936	43.936	43.936	43.936
Production	7.559	71.396	43.223	33.760
Production	42.355	42.355	42.355	42.355
Production	15.379	60.084	38.211	31.363
Production	22.413	53.173	37.793	37.793
Production	35.908	35.908	35.908	35.908
Production	12.962	88.555	35.785	21.586
Production	22.783	46.247	35.180	36.511
Production	34.861	34.861	34.861	34.861
Production	13.585	53.251	33.418	33.418
Production	11.614	69.887	33.394	18.679
Production	29.791	33.927	31.859	31.859
Production	13.179	53.955	31.255	26.632
Production	23.332	37.791	30.562	30.562
Production	18.483	39.220	28.851	28.851
Production	27.880	27.880	27.880	27.880
Production	2.259	37.446	27.553	29.310
Production	24.590	24.590	24.590	24.590
Production	23.246	23.246	23.246	23.246
Production	4.213	39.737	21.975	21.975
Production	21.368	21.368	21.368	21.368
Production	21.112	21.112	21.112	21.112
Production	6.995	33.154	20.075	20.075
Production	19.233	20.779	20.006	20.006
Production	19.891	19.891	19.891	19.891
Production	1.964	35.942	18.953	18.953
Production	1.187	36.597	18.892	18.892
Production	3.918	32.907	18.413	18.413
Production	18.085	18.085	18.085	18.085
Production	17.550	17.550	17.550	17.550
Production	17.383	17.383	17.383	17.383
Production	9.874	27.032	17.150	17.029
Production	10.498	20.457	17.137	20.457
Production	0.702	30.464	15.788	16.198
Production	15.689	15.689	15.689	15.689
Production	14.914	14.914	14.914	14.914
Production	14.777	14.777	14.777	14.777

Production	14.240	14.240	14.240	14.240
Production	4.477	23.818	14.148	14.148
Production	5.705	21.726	13.715	13.715
Production	13.143	13.143	13.143	13.143
Production	4.509	27.957	12.535	11.273
Production	6.619	16.983	11.801	11.801
Production	11.685	11.685	11.685	11.685
Production	11.625	11.625	11.625	11.625
Production	2.779	19.885	11.332	11.332
Production	9.095	13.250	11.172	11.172
Production	10.431	10.431	10.431	10.431
Production	10.063	10.063	10.063	10.063
Production	9.503	9.503	9.503	9.503
Production	2.892	15.974	9.433	9.433
Production	3.058	15.400	9.229	9.229
Production	1.883	15.991	8.937	8.937
Production	3.607	13.434	8.521	8.521
Production	3.758	13.267	8.512	8.512
Production	3.971	11.361	7.666	7.666
Production	7.633	7.633	7.633	7.633
Production	7.029	7.029	7.029	7.029
Production	6.551	6.551	6.551	6.551
Production	2.223	9.339	5.781	5.781
Production	3.304	7.469	5.387	5.387
Production	3.856	6.415	5.136	5.136
Production	5.044	5.044	5.044	5.044
Production	5.037	5.037	5.037	5.037
Production	4.563	4.997	4.780	4.780
Production	4.238	4.238	4.238	4.238
Production	3.909	3.909	3.909	3.909
Production	1.825	5.969	3.897	3.897
Production	0.957	6.535	3.746	3.746
Production	3.488	3.488	3.488	3.488
Production	3.409	3.409	3.409	3.409
Production	3.237	3.237	3.237	3.237
Production	2.671	2.671	2.671	2.671
Production	2.444	2.444	2.444	2.444
Production	1.767	1.767	1.767	1.767
Production	0.875	0.875	0.875	0.875

Table 20: US EUR data

Region	EUR (Mm^3)
US	8.494
US	8.494
US	11.325
US	14.000
US	14.156
US	15.501
US	15.572
US	25.481
US	28.313
US	28.313
US	28.313
US	33.975
US	34.000
US	40.204
Marcellus	43.885
US	45.000
US	56.625
US	56.625
US	56.625
US	57.000
US	62.288
US	73.613
US	76.444
US	77.577
Barnett	84.938
US	84.938
US	84.938
US	85.000
US	85.000
US	99.094
US	150.000
US	150.000
US	198.189
US	2576.456
US	2583.534

Table 21: Input data for Clearstone and GreenPath's top 5% emitters in emission source population

Reporting Segment	Emission Type	Emission Source	Emission Volume (m^3/day)
Compressor Station	Leak	Reciprocating Compressor	57.572
Gas Gathering System	Leak	Reciprocating Compressor	80.900
Gas Gathering System	Leak	Reciprocating Compressor	57.262
Gas Multiwell Group	Vent	Separator	8.042
Gas Single	Leak	Separator	10.462
Gas Multiwell Group	Leak	Wellhead	63.272
Gas Multiwell Group	Leak	Wellhead	75.828
Gas Multiwell Group	Leak	Wellhead	1126.646
Gas Gathering System	Leak	Dehydrator - Glycol	4.447
Gas Gathering System	Unlit Flare	Flare KnockOut Drum	84.214
Compressor Station	Vent	Production Tank (fixed roof)	360.216
Gas Multiwell proration outside SE AB	Leak	Gas Pipeline Header	6.885
Gas Gathering System	Leak	Screw Compressor	11.558
Gas Gathering System	Leak	Electric Driver Compressor	15.446
Gas Well	Leak	Wellhead	20.341
Gas Well	Leak	Separator	25.427

Table 22: Input data for Clearstone and GreenPath's top 5% emitters in reporting segment population

Reporting Segment	Emission Type	Emission Source	Emission Volume (m^3/day)
Gas Battery	Vent	Production Tank (fixed roof)	34.725
Gas Battery	Leak	Wellhead	63.272
Gas Battery	Leak	Wellhead	75.828
Gas Battery	Leak	Wellhead	41.933
Gas Battery	Leak	Wellhead	1126.646
Gas Battery	Leak	Reciprocating Compressor	52.980
Compressor Station	Vent	Production Tank (fixed roof)	360.216
Gas Gathering System	Leak	Reciprocating Compressor	80.900
Gas Gathering System	Unlit Flare	Flare KnockOut Drum	84.214
Gas Gathering System	Leak	Production Tank (fixed roof)	165.726
Gas Gathering System	Leak	Reciprocating Compressor	57.262
Gas Battery	Leak	Tankage	1525.608
Gas Well	Leak	Filter/Separation	25.427

Table 23: Input data for Zavala-Araiza et al (Red Deer) and O'Connell et al (Medicine Hat)'s top 5% emitters

Region	Emission Volume (m^3/day)
Red Deer	2399.700
Red Deer	518.100
Red Deer	639.000
Medicine Hat	820.170
Medicine Hat	290.903
Medicine Hat	248.762
Medicine Hat	199.226
Medicine Hat	107.891