Graduate Theses and Dissertations

5-2016

Quantifying, Assessing, and Mitigating Methane Emissions from Super-emitters in the Oil and Gas Supply Chain

David Richard Lyon
University of Arkansas, Fayetteville

Quantifying, Assessing, and Mitigating Methane Emissions
from Super-emitters in the Oil and Gas Supply Chain

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Environmental Dynamics

by

David Richard Lyon
Hendrix College
Bachelor of Arts in Biology, 2002
University of Kentucky
Master of Science in Forestry, 2004

May 2016
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

_____
Dr. Phillip Hays
Dissertation Director

_____   _____
Dr. Steven Boss   Dr. Ralph Davis
Committee Member   Committee Member

_____
Dr. Benjamin Runkle
Committee Member

ABSTRACT

Methane emissions from the oil and gas (O&G) supply chain reduce potential climate benefits of natural gas as a replacement for other fossil fuels that emit more carbon dioxide per energy produced. O&G facilities have skewed emission rate distributions with a small fraction of sites contributing the majority of emissions. Knowledge of the identity and cause of these high emission facilities, referred to as super-emitters or fat-tail sources, is critical for reducing supply chain emissions. This dissertation addresses the quantification of super-emitter emissions, assessment of their prevalence and relationship to site characteristics, and mitigation with continuous leak detection systems. Chapter 1 summarizes the state of the knowledge of O&G methane emissions. Chapter 2 constructs a spatially-resolved emission inventory to estimate total and O&G methane emissions in the Barnett Shale as part of a coordinated research campaign using multiple top-down and bottom-up methods to quantify emissions. The emission inventory accounts for super-emitters with two-phase Monte Carlo simulations that combine site measurements collected with two approaches: unbiased sampling and targeted sampling of super-emitters. More comprehensive activity data and the inclusion of super-emitters, which account for 19% of O&G emissions, produces a emission inventory that is not statistically different than top-down regional emission estimates. Chapter 3 describes a helicopter-based survey of over 8,000 well pads in seven basins with infrared optical gas imaging to assess high emission sources. Four percent of sites are observed to have high emissions with over 90% of observed sources from tanks. The occurrence of high emissions is weakly correlated to site parameters and the best statistical model explains only 14% of variance, which demonstrates that the occurrence of super-emitters is primarily stochastic. Chapter 4 presents a Gaussian dispersion model for optimizing the placement of continuous leak detection systems at three example well pads. The model demonstrates that large leaks can be detected quickly with first generation systems. Continuous leak detection can be used in the near future to cost-effectively mitigate methane emissions from O&G super-emitters.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF PUBLISHED PAPERS

Chapter 1 (published)

Lyon, D.R. 2016. Methane emissions from the natural gas supply chain. In: Kaden, D.A. and Rose, T.L. eds. *Environmental and Health Issues in Unconventional Oil and Gas Development*. Elsevier. pp. 33-48.

Chapter 2 (published)

Lyon, D.R., Zavala-Araiza, D., Alvarez, R.A., Harriss, R., Palacios, V., Lan, X., Talbot, R., Lavoie, T., Shepson, P., Yacovitch, T.I. Herndon, S.C., Marchese, A.J., Zimmerle., D., Robinson, A.L. and Hamburg, S.P. 2015. Constructing a spatially resolved methane emission inventory for the Barnett Shale region. *Environmental Science & Technology*. 49(13), pp.8147-8157. DOI: 10.1021/es506359c

Chapter 3 (published)

Lyon, D.R., Alvarez, R.A., Zavala-Araiza, D., Brandt, A.R., Jackson, R.B. and Hamburg, S.P., 2016. Aerial surveys of elevated hydrocarbon emissions from oil and gas production sites. *Environmental Science & Technology*. Article ASAP. DOI: 10.1021/acs.est.6b00705

INTRODUCTION

Emissions of methane, a short-lived climate pollutant and the primary constituent of natural gas, can diminish or delay climate benefits of substituting natural gas for other fossil fuels. Numerous studies have found that oil and gas (O&G) facilities have highly skewed emission rate distributions with a small number of facilities responsible for the majority of total emissions. The lack of knowledge about these high emission facilities, often referred to as super-emitters, has contributed to uncertainty in the quantification of O&G supply chain emissions and hindered the formulation of effective policies for reducing emissions. This dissertation is divided into four chapters focused on O&G super-emitters. Chapter 1 summarizes the state of the knowledge on methane emissions from the O&G supply chain. Chapter 2 quantifies super-emitters to more accurately estimate regional O&G methane emissions. Chapter 3 assesses the prevalence and relationship of super-emitters to site characteristics to determine if their occurrence can be predicted. Chapter 4 describes a dispersion model used to assess the performance of continuous leak detection systems for mitigating super-emitter emissions.

Chapter 1, *Methane emissions from the natural gas supply chain*, presents background information on the importance of methane emissions to climate change and the various sources of methane from the O&G supply chain. A literature review describes the state of knowledge prior to unconventional development and summarizes several recent research studies. The chapter introduces the concept of super-emitters and their importance in reducing supply chain emissions.

Chapter 2, *Constructing a spatially resolved methane emission inventory for the Barnett Shale Region*, addresses the observed discrepancy between top-down and bottom-up estimates of O&G methane emissions. The chapter describes the development a detailed emission inventory that integrates spatially-referenced activity data with emission factors based on recent measurement data from a coordinated research campaign in the Barnett Shale that used multiple top-down and bottom-up approaches to quantify methane emissions. Site measurements were collected by other research teams with two types of sampling approaches: unbiased and high-biased sampling. Unbiased sampling is necessary to characterize the lower and middle portions of site emission rate distributions, but typically will underestimate average emissions of the population since a relatively small sample size is unlikely to

include super-emitters from the upper portion of the distribution. Targeted sampling of high emission sites can quantify emissions from super-emitters but will overestimate emissions if they are overrepresented in emission factors. Therefore, two-phase Monte Carlo simulations are used to blend unbiased and high-biased measurements into single emission factors for well pads, small processing plants, large processing plants, and compressor stations. The emission inventory is compared to top-down measurements of regional emissions and alternative bottom-up estimates based on the U.S. EPA Greenhouse Gas Inventory and Greenhouse Gas Reporting Program.

Chapter 3, *Aerial surveys of elevated hydrocarbon emissions from oil and gas production sites*, evaluates the prevalence of super-emitters at well pads across the U.S. and attempts to predict their occurrence with statistical models. Over 8,000 well pads in seven basins are selected by stratified random sampling to produce a sample covering over 1% of U.S. active wells and the diversity of production types. A professional leak detection firm is contracted to survey these sites with a helicopter-based infrared camera that makes large hydrocarbon emission sources visible. The survey detection limit is estimated with a controlled release and aircraft-based quantification of methane emissions at a subset of surveyed sites. A national well database is used to determine the well count, well age, gas production, oil production, water production, and percent energy from oil at surveyed sites. Single parameter and multi-parameter generalized linear models using basin, site, and operator parameters are evaluated to explain the variance in observed high emissions. Finally, a tank flashing analysis is performed to assess if the observed prevalence can be explained by potential or controlled tank flashing emissions.

Chapter 4, *Site-level Gaussian dispersion model to optimize the deployment of continuous methane sensors*, describes a model programmed by the author in the open source language R to assess the performance of continuous leak detection systems at O&G sites. The model uses Gaussian dispersion equations, local meteorological data, and the location of potential emission sources to estimate the methane concentration enhancement caused by individual sources at multiple receptors. Based on sensor detection limits in resolving concentration enhancements from background, the model calculates the average and 95th percentile hours to detect different emission rates by both point and open-path sensors deployed at various receptors. The model is used to assess the performance of sensors with

three different detection limits for detecting large and small emission rates at three example well pads in the Eagle Ford, Fayetteville, and Bakken Shale. The results provide information on the ideal locations to deploy sensors to most quickly detect emissions.

CHAPTER 1

Methane emissions from the natural gas supply chain

Introduction

Natural gas production in the United States increased 36% between 2005 and 2014, primarily due to the development of shale gas resources with horizontal drilling and hydraulic fracturing (EIA 2015). Increased gas supply led to lower gas prices in the U.S., which contributed to a 21% decrease in coal-fired electricity generation over this time period (EIA 2015). The replacement of coal with natural gas for electricity generation has long-term climate benefits since natural gas combustion produces less carbon dioxide (CO_2) per unit of energy generated than coal. However, methane (CH_4), the primary constituent of natural gas, is a powerful greenhouse gas with 120 times the radiative forcing of CO_2 on a mass basis (IPCC 2013). Methane has an effective atmospheric lifetime of approximately 12 years while a large portion of CO_2 emissions persists in the atmosphere for much longer. Therefore, the climate impact of methane is greater in the short-term with a 20-year global warming potential (GWP) of 84- and a 100-year GWP of 25; with the inclusion of climate-carbon feedbacks, GWP values increase to 86 and 32, respectively (IPCC 2013).

Alvarez et al. (2012) developed the concept of technology warming potential to compare the relative climate impacts over time of different technologies. The climate impacts of natural gas-fueled technologies are dependent on methane emissions across the natural gas supply chain. As shown in Figure 1, replacement of a coal-fired power plant with a gas-fired plant will have immediate climate benefits if the supply chain methane emissions are less than 2.7% of produced natural gas (Alvarez et al. 2012; updated with IPCC 2013). At higher gas leak rates, switching from more carbon-intensive fossil fuels to gas will be worse for the climate in the near term due to the short-term climate impacts of

4

methane. Although CO_2 emission reductions are necessary for long-term mitigation, reducing emissions of short-lived climate pollutants, such as methane, has the advantage of potentially delaying climate system tipping points and allowing more time for adaptation (Shoemaker et al. 2013). Consequently, it is critical to quantify and mitigate methane emissions from the natural gas supply chain to minimize its climate impacts and assure that gas displacement of other fossil fuels has immediate climate benefits.

Emission Sources

Methane emissions occur across the entire natural gas supply chain. The production sector includes the short-term development of new wells and the long-term operation of producing wells. Since many unconventional wells co-produce oil and natural gas, this chapter will consider oil and gas production as a single sector, but focus on natural gas in downstream sectors. The gathering sector is a system of pipelines and compressor stations that transports gas from well pads to processing plants or transmission pipelines. The processing sector includes plants that treat gas to pipeline-quality standards, which can include removal of water, CO_2, hydrogen sulfide, and natural gas liquids (heavier hydrocarbons, such as ethane). Some produced gas is close to pipeline-quality standards and may bypass the processing sector and instead undergo minor treatment, such as dehydration, in the production or gathering sectors. The transmission and storage sector is a system of high-pressure pipelines and compressor stations that transports gas from gathering systems and processing plants to high-demand customers such as power plants and local distribution city gates. This sector also includes underground and liquefied natural gas (LNG) storage facilities that store gas for periods of higher demand. Finally, the local distribution sector is a system of pipelines and metering and regulating stations that delivers gas from city gates to customers, such as commercial and residential buildings.

Methane emissions from the natural gas supply chain can be grouped by three basic source classifications: vented emissions, fugitive emissions, and incomplete combustion emissions. Vented emissions are intentional releases related to normal operations or safety procedures. Fugitive emissions are unintentional releases from equipment leaks and malfunctioning equipment. Incomplete combustion emissions are fuel slippage in exhaust of natural gas combustion sources.

Natural gas-powered pneumatic controllers are the largest source of vented emissions in the U.S. These devices use the energy of pressurized gas to power equipment controlling process variables such as liquid level and temperature (Simpson 2014). Depending on their design, pneumatic controllers emit gas continuously or intermittently during normal operation. Similarly, pneumatic pumps use pressurized gas to power chemical injection or glycol pumps.

Intentional release of gas for safety or convenience is another large source of vented emissions. Blowdowns occur when equipment is depressurized prior to maintenance or to alleviate over-pressurization. Associated gas venting is the release of gas produced by oil wells that lack a connection to gas gathering systems. Some wells accumulate liquids in the wellbore that inhibit gas production; these wells may be temporarily vented to atmosphere to unload liquids, which also emits methane.

Well completions are a vented emission source specific to unconventional development. After a well is hydraulically fractured, excess fluid and proppant must be cleared from the wellbore. In the earlier years of unconventional development, completions were vented to the atmosphere, which released methane in flowback gas. As of January 2015, a United States Environmental Protection Agency (U.S. EPA) rule requires almost all hydraulically-fractured gas wells, but not oil wells, to use reduced emission completions to capture completion flowback gas for sales instead of venting the gas to the atmosphere ; a proposed rule would extend this requirement to oil wells (40 CFR Part 60 Subpart OOOO).

Hydrocarbon and produced water storage tanks have vented emissions referred to as flashing losses, working losses, and breathing losses. After fluid flows from a high-pressure separator to an atmospheric pressure tank, entrained gas is emitted as flashing losses. Working losses occur from displacement of vapors when a tank is filled. Breathing losses are caused by changes in ambient temperature or pressure. Similar to storage tank flashing, glycol dehydrators can emit methane entrained in wet glycol through the vent that releases water vapor. Tank emissions can be controlled with flares or vapor recovery units that combust or capture these losses.

Equipment leaks are fugitive emissions from poorly sealed or damaged components that are designed to have zero emissions, such as connectors and pipelines. Fugitive emissions can also occur from

components that are expected to have some leakage — for instance, the vents of reciprocating compressor rod packing seals. Although these sources are designed to vent some gas during normal operation, their emission rates can increase with equipment wear. Malfunctions of venting sources or control equipment can cause fugitive emissions. For example, a malfunctioning pneumatic controller can emit beyond its designed rate, or tank emissions can bypass a flare if the tank hatch is open.

Incomplete combustion emissions occur from engines, flares, or other equipment such as heaters that combust natural gas. Reciprocating compressor engines normally emit a greater portion of their fuel throughput than centrifugal compressors (EPA 1995). Natural gas flares are often assumed to combust 98% of gas, but a recent study reported >99.8% methane combustion efficiency for 11 flares in North Dakota and Pennsylvania (Caulton et al. 2014).

State of Knowledge Prior to Unconventional Development

The U.S. EPA annually publishes two sources with methane emission estimates from the U.S. natural gas supply chain: the U.S. Greenhouse Gas Inventory (GHGI; EPA 2015a) and the Greenhouse Gas Reporting Program (GHGRP; EPA 2015b). Both these sources rely heavily on data collected during a comprehensive study in the early 1990s by the U.S. EPA and the Gas Research Institute that estimated 1992 U.S. natural gas supply chain methane emissions were equivalent to 1.4±0.5% of gross natural gas production (Harrison et al. 1996).

The GHGI is an annual report that includes estimates of U.S. annual greenhouse gas emissions by source category from 1990 to two years prior of the publication year; it is published in fulfillment of the U.S.'s commitments under the United Nations Framework Convention on Climate Change. Emissions are reported at the national level for natural gas systems and petroleum systems except for the gas production sector, which is reported at the regional level. The GHGI estimates activity factors (equipment and facility counts) based on a combination of recent data, such as well counts, and assumptions, such as the number of pneumatic controllers per well, based on Harrison et al. (1996) and other studies. For most source categories, potential emissions are estimated by multiplying activity factors by emission factors (average emissions per equipment or site) based on Harrison et al. (1996) and other studies. Net

emissions are calculated by subtracting estimated emission reductions, which are due to regulations and the voluntary U.S. EPA Natural Gas STAR program, from potential emissions. Each year, U.S. EPA updates the GHGI methodology and recalculates annual emissions back to 1990.

The GHGI has undergone several methodological changes in recent years that greatly changed estimates of natural gas systems methane emissions. Most recent year emission estimates increased from 4.6 teragrams (Tg) methane in the 2010 GHGI to 10.5 Tg in the 2011 GHGI, primarily due to the addition of a source category for hydraulically fractured well completions and a methodological change resulting in increased emission estimates from liquids unloading. The 2013 GHGI emission estimate was reduced to 6.9 Tg due to another revision in liquids unloading methodology. A revised methodology for hydraulically fractured well completions reduced the 2014 GHGI estimate to 6.2 Tg. The 2015 GHGI estimates 2013 emissions from natural gas systems are 6.3 Tg with an additional 1.0 Tg from petroleum systems production. Assuming a supply chain average gas composition of 90% methane, 7.3 Tg methane is equivalent to 1.4% of gross natural gas production.

The GHGRP is a mandatory reporting program for U.S. facilities with annual GHG emissions ≥25,000 metric tons carbon dioxide equivalents (CO_2e). Petroleum and natural gas facilities report under Subpart W, which includes onshore petroleum and gas production, offshore petroleum and gas production, gas processing, gas transmission, underground gas storage, LNG storage, LNG import/export, and gas local distribution (40 CFR Part 98). These facilities also report incomplete combustion emissions under Subpart C (40 CFR Part 98) Gas gathering and boosting facilities report under Subpart W but the U.S. EPA has amended the rule to also require reporting under Subpart W beginning in 2016. Facilities are defined at the site-level except for onshore production, which are defined as a company's entire well pad assets in a basin, and local distribution, which are defined as a company's statewide distribution assets. Since 2011, reporters have been required to report annual emissions by source category and associated data at levels of detail varying by source and sector. Emissions must be estimated using methods prescribed by the rule for each source category, which includes direct measurements, engineering equations, and emission factors. Similar to the GHGI, GHGRP emission factors are based primarily on Harrison et al.

(1996). The GHGRP should not be viewed as a comprehensive inventory since it only includes facilities above the reporting threshold and excludes the gathering sector and some emission sources.

Recent Studies

Research on natural gas supply chain methane emissions has resurged over the last several years due to concern regarding the climate impacts of increased natural gas development, particularly since a 2011 paper estimated that shale gas has a 3.6 – 7.9% leak rate (Howarth, Santoro, & Ingraffea 2011). U.S. EPA GHGI and GHGRP emission estimates depend on data from Harrison et al. (1996) and other studies collected prior to the growth in unconventional oil and gas (O&G) development. Numerous recent studies have investigated whether unconventional development or other changes in operational practices have increased or decreased methane emissions, including a series of sixteen studies sponsored by Environmental Defense Fund. Research can be divided into two categories based on their methodological approaches: top-down studies that use atmospheric measurements of well-mixed air to estimate emissions at a regional or larger scale, and bottom-up studies that measure emissions at the component or site-level and sometimes extrapolate emissions to a regional or national scale using activity and emission factors. Bottom-up studies are summarized by sector below, followed by top-down studies of total O&G emissions.

Production

Allen et al. (2013) used direct measurements to quantify emissions from equipment leaks, pneumatic controllers, and pneumatic pumps at 150 onshore production sites, 27 well completion flowbacks, 4 well workovers, and 9 liquid unloading wells located across the U.S. Compared to GHGI estimates, emissions were higher for pneumatic controllers and equipment leaks, but lower for well completions. Two-thirds of well completions sent flowback gas to a sales line or flare, which controlled 99% of potential emissions. This demonstrates that well completion flowbacks, the main difference between conventional and unconventional production, can be effectively controlled with existing technologies.

Two follow up studies to Allen et al. (2013) measured emissions from 107 liquid unloading wells and 377 pneumatic controllers (Allen et al. 2015a, Allen et al. 2015b). Liquids unloading emissions were similar to

the most recent GHGI estimates with the highest emissions from wells with more than 100 annual unloading events. Pneumatic controllers had 17% higher average emissions per device than the GHGI, and the average number of controllers per well was 2.7 compared to 1.0 in the GHGI, which indicates that the GHGI underestimates their national emissions by up to a factor of 3. A small number of controllers, many that were malfunctioning, had emission rates exceeding 6 standard cubic feet per hour (scfh). These devices comprised 19% of the population but accounted for 95% of emissions.

Allen et al. (2013) estimated emissions from the U.S. natural gas production sector by scaling up measured sources using national activity factors and assuming GHGI estimates were accurate for unmeasured sources. National emission estimates were updated in the follow up studies (Allen et al. 2015a, Allen et al. 2015b) using 2012 activity factors and new pneumatic controller and liquids unloading data. Production sector 2012 emissions were estimated to be 2,185 gigagrams (Gg) methane, equivalent to 0.49% of gross gas production. The upper bound emission estimate was 2,815 Gg due to uncertainty in the pneumatic controller and liquids unloading emissions. In comparison, GHGI and GHGRP 2012 production emissions are 1,992 Gg and 2,200 Gg, respectively; the GHGI and Allen et al. estimates are for natural gas only, while the GHGRP includes oil and gas production.

Brantley et al. (2014) applied U.S. EPA Other Test Method (OTM) 33A, a mobile sampling method using point source Gaussian dispersion modeling, to estimate emissions at 210 production sites in the Barnett, Denver-Julesburg, Pinedale, and Eagle Ford basins. Emission rates were log-normally distributed and had a weak positive correlation (R^2 = 0.083) with gas production. Compared to Allen et al. (2013) and a study in the Barnett Shale (ERG 2011), emission rate geometric means were about twice as high in Brantley et al. (2014), which may be due to the exclusion of tank flashing emissions in these other studies or the bias of mobile sampling towards higher emitting sites with detectable downwind plumes.

Gathering & Processing

A recent study measured site-level emissions at 114 U.S. gathering facilities and 16 processing plants using the dual tracer correlation method (Roscioli et al. 2015, Mitchell et al. 2015). Gathering facility emissions were positively skewed with 12% of sites contributing 50% of emissions. Infrared camera

surveys revealed that 20% of gathering sites had venting tanks and four times the average emissions of sites without substantial venting. Processing plants had higher average emission rates than gathering facilities (170 vs. 55 kg methane h^{-1}) but lower emissions proportional to gas throughput (0.075% vs. 0.2%). For both facility types, throughput was positively correlated with absolute emissions and negatively correlated with throughput-normalized emissions.

Marchese et al. (2015) estimated emissions from the U.S. gathering and processing sectors using Monte Carlo simulations incorporating facility counts and emissions data from Mitchell et al. (2015). The number of U.S. processing plants (606) is obtainable from national databases, but there is no comprehensive, national list of gathering facilities. The authors estimated the number of gathering facilities, 4,549 (+921/-703), by comparing state permit databases to lists obtained from study industry participants. Gathering and processing (G&P) sector 2012 emissions were estimated to be 2,421 (+245/-237) Gg methane compared to 1,296 Gg in the 2014 GHGI and 180 Gg in the 2013 GHGRP. Processing sector emissions were estimated to be 546 Gg, lower than the GHGI (892 Gg), but higher than the GHGRP (179 Gg). Since the GHGI includes gathering within the production sector, the authors used industry participant facility equipment counts to allocate GHGI emissions between production and gathering. The study estimate of gathering sector emissions (1,875 Gg) is almost 5 times the GHGI estimate (404 Gg) and several orders of magnitude higher than the GHGRP (0.5 Gg), which excluded vented and fugitive emissions from gathering facility reporting requirements. G&P sector emissions are equivalent to 0.47±0.05% of gas production.

Transmission & Storage

Subramanian et al. (2015) quantified methane emissions at 45 transmission and storage (T&S) sector compressor stations using component-level direct measurements and the site-level dual tracer correlation method. Study onsite estimates used infrared camera surveys to identify emitting sources followed by quantification with high-flow samplers; incomplete combustion emissions were estimated with U.S. EPA AP-42 emission factors (EPA 1995). The average tracer flux emission rate was 80 kg methane h^{-1}. Similar to the G&P sector, T&S sites had a skewed emission rate distribution with 50% of emissions from 10% of sites. The two highest emitting sites, which were defined as super-emitters, had much higher site-level

emissions based on the tracer correlation method than aggregate measured component-level emissions. This discrepancy was caused by the presence of leaking compressor isolation valves that could not be accurately quantified with component-level measurements. For the 25 sites exceeding the GHGRP reporting threshold, study onsite emissions were 1.8 times higher than emissions estimated using prescribed GHGRP methods. This difference is due to GHGRP methodologies that exclude certain sources (e.g., reciprocating compressor rod-packing vents in pressurized, standby mode) and require use of inaccurate emission factors, including an incomplete combustion factor that is over 500 times lower than the AP-42 factor for reciprocating compressors.

Emissions from the U.S. T&S sector were estimated in Zimmerle et al. (2015) with Monte Carlo simulations that integrated emissions data from Subramanian et al. (2015), detailed facility data from study industry participants, and GHGRP data. National emission estimates for 2012 were 1,503 (+750/-283) Gg methane compared to GHGI and GHGRP estimates of 2,071 and 200 Gg methane, respectively. In contrast to the GHGI, Zimmerle et al. (2015) estimated that there are fewer T&S stations in the U.S. and the compressors at these sites have a greater proportion of lower emitting dry seal centrifugal compressors. T&S sector emissions are equivalent to 0.35% (+0.10/-0.07%) of T&S sector throughput.

Local Distribution

Lamb et al. (2015) measured methane emissions at 230 pipeline leaks and 229 metering and regulating (M&R) stations at 13 U.S. local distribution systems. Sites were randomly selected from lists of pipeline leaks and M&R stations in representative areas provided by study industry participants. High-flow samplers were used to directly measure M&R components and in tandem with surface enclosures to measure underground pipeline leaks. There was an extremely skewed emission rate distribution for pipeline leaks with 1.3% of leaks contributing 50% of total emissions. Compared to Harrison et al. (1996), which is the basis of the GHGI emission factors, both pipeline leaks and M&R stations had lower average emission rates. Nine M&R stations were measured in both Lamb et al. (2015) and Harrison et al. (1996), eight of which had lower emissions in the more recent study. Lamb et al. (2015) estimated 2011 U.S. local distribution emissions using the new measurement data. Emission factors were developed for pipeline mains and services by pipeline material and M&R stations by operating pressure, and then

multiplied by GHGI activity factors. GHGI emission estimates were used for customer meter and upset emissions. Local distribution 2011 emissions were estimated to be 393 Gg methane (95% upper confidence limit = 854 Gg), compared to 1,329 and 640 Gg in the GHGI and GHGRP respectively. In contrast to the GHGI, study estimates were about 3 times lower for pipeline leaks and 13 times lower for M&R stations. The substantially lower emissions were attributed to pipeline replacement, leak surveys, and station upgrades and maintenance since the 1990s.

Top-down Studies

Many recent studies have used aircraft, tower, or satellite-based measurements to estimate total methane emissions in regions with O&G production and/or natural gas distribution (Petron et al. 2012, Townsend-Small et al. 2012, Jeong et al. 2013, Karion et al. 2013, Miller et al. 2013, Kort et al. 2014, Petron et al. 2014, Wecht et al. 2014, Peischl et al. 2015). Some of these studies have estimated the fraction of emissions from O&G sources using source apportionment approaches including stable isotope and hydrocarbon ratios, or by subtracting bottom-up estimates of other sources (Townsend-Small et al. 2012, Petron et al. 2014). There is large variability among basins in top-down estimates of methane emissions as a percentage of gas production. For example, aircraft mass balance studies have reported leak rates of 0.18 - 0.41% in the Marcellus, 1.0 - 2.1% in the Haynesville, 1.0 - 2.8% in the Fayetteville, 2.6 – 5.6% in the Denver-Julesburg, and 6.2 – 11.7% in the Uintah (Peischl et al. 2015, Petron et al. 2014, Karion et al. 2013).

Top-down estimates of total and O&G methane emissions often have been higher than bottom-up estimates. Miller et al. (2013) analyzed methane observations with an atmospheric transport model and geostatistical inverse modeling to estimate U.S. emissions. Anthropogenic emission estimates were 1.5 and 1.7 times higher than the GHGI and the Emissions Database for Global Atmospheric Research (EDGAR; JCR/PBL 2011), respectively; O&G emissions in the south-central U.S. were estimated to be 2.3 – 7.5 times higher than EDGAR. A review of research studies found that top-down estimates typically exceed bottom-up estimates; the authors estimate U.S. methane emissions are 1.25 – 1.75 times higher than GHGI estimates (Brandt et al. 2014). The authors propose four hypotheses that may account for underestimation by O&G emission inventories: 1) bottom-up data are no longer representative of current

technologies and practices, 2) bottom-up datasets have small sample size and possible sampling bias, 3) skewed emission distributions result in sampled data with lower average emission rates than the population, and 4) inaccurate activity factors.

In October 2013, a coordinated research campaign in the Barnett Shale used simultaneous top-down and bottom-up approaches to quantify methane emissions from the O&G supply chain (Harriss et al. 2015). Bottom-up data included direct component measurements of T&S stations (Johnson et al. 2015) and local distribution M&R stations (Lamb et al. 2015). Aircraft- and vehicle-based approaches were used to quantify site-level emissions at O&G well pads, compressor stations, processing plants, and landfills by analyzing downwind plumes with the mass balance approach (Lavoie et al. 2015, Nathan et al. 2015), Gaussian dispersion modeling (Lan et al. 2015, Yacovitch et al. 2015), or mobile flux plane method (Rella et al. 2015). Regional methane emissions were estimated using the top-down, aircraft-based mass balance approach during 8 flights (Karion et al. 2015). Source apportionment data included carbon and hydrogen stable isotope and hydrocarbon ratios of canister air samples from different source types (Townsend-Small et al. 2015) and continuous ethane-to-methane ratios measured on the mass balance aircraft (Smith et al. 2015). Top-down estimates of total methane emissions were 76±13 Mg h^{-1} with 60±11 Mg h^{-1} from fossil sources (Karion et al. 2015, Smith et al. 2015). Lyon et al. (2015) used data from the campaign and other sources to construct a spatially-resolved methane emission inventory; bottom-up estimates of total and O&G methane emissions were 72.3 (+10.1/-8.9) Mg h^{-1} and 46.2 (+8.2/-6.2) Mg h^{-1}, respectively. This bottom-up estimate of O&G emissions was 1.5, 2.7, and 4.3 times higher than alternative estimates based on the GHGI, GHGRP, and EDGAR, respectively, primarily due to the inclusion of high-emission sites and more complete activity factors, particularly for gathering stations. Top-down and bottom-up estimates of the Barnett region natural gas supply chain leak rate (1.3 – 1.9% and 1.0 – 1.4%, respectively) are not significantly different. Zavala-Araiza et al. (2015) used well pad data from the campaign to introduce the concept of functional superemitters defined by emissions proportional to gas production. Functional superemitters comprised 15% of sites and contributed 77% of total production site emissions in the Barnett region.

Remaining Uncertainties

Compared to the GHGI, bottom-up studies have reported similar emissions for production and T&S (Allen et al. 2015a, Zimmerle et al. 2015), higher emissions for gathering (Marchese et al. 2015), and lower emissions for processing and local distribution (Marchese et al. 2015, Lamb et al. 2015). A forthcoming study will integrate recent data sources to estimate the magnitude and uncertainty of U.S. natural gas supply chain emissions. There are several sources of remaining uncertainty that may affect these emission estimates.

Positively skewed emission rate distributions are common to many types of O&G facilities and components. As shown in Figure 2, a relatively small number of high emission sites, sometimes referred to as superemitters, contribute a large fraction of emissions. Recent studies have used statistical methods to quantify the uncertainty associated with skewed distributions; for example, the upper confidence limit of U.S. local distribution emissions was over twice the central estimate mainly due to the majority of measured pipeline emissions coming from a very small number of leaks (Lamb et al. 2015). Additional research on the prevalence, magnitude, and causes of super-emitters may reduce the uncertainty associated with sampling skewed distributions.

Activity factors for some equipment and facilities are poorly known. Allen et al. (2015a) reported that the average number of pneumatic controllers per well was 2.7 times higher than GHGI estimates. The number of gathering facilities is especially uncertain since these sites are subject to less reporting requirements. Marchese et al. (2015) improved estimates of U.S. gathering facility counts but their lower and upper bound estimates still vary by a factor of 1.4. More comprehensive reporting of activity factors is critical for reducing uncertainty in bottom-up emission estimates.

Several emission sources that were not measured by recent studies continue to have uncertainty associated with their previous emission estimates. Production sector storage tanks and compressors were not measured fully by Allen et al. (2013). Other studies have indicated that high emission well pads and gathering stations are often associated with tank venting (Brantley et al. 2014, Mitchell et al. 2015.), which supports the need for additional data on tank emissions. Gathering pipelines are not known to have any published emissions data. The GHGI uses data from Harrison et al. (1996) that was based on local distribution main leaks to estimate gathering pipeline emissions; Marchese et al. (2015) used the GHGI

data in their updated estimate of gathering sector emissions. Gathering pipelines, which are less regulated than either transmission or local distribution pipelines, are potentially a large emission source that should be a focus of future research.

Emissions may also be associated with natural gas end use past the customer meter, such as leaks at power plants and incomplete combustion by residential gas furnaces. A top-down study estimated Boston region emissions are 2.7±0.6% of gas throughput compared to a bottom-up estimate of 1.1% (McKain et al. 2014); the authors hypothesize that the gap may be partially due to end use emissions. A forthcoming study will report emissions data from natural gas vehicles and fueling stations (Clark et al., in review).

Source apportionment methods such as stable isotope ratios cannot distinguish methane emissions from the natural gas supply chain and other fossil sources. Emissions from abandoned O&G wells and geologic seepage may be responsible for some of the gap between top-down and bottom-up estimates (Brandt et al. 2014). There are approximately 3 million abandoned and inactive wells in the U.S., but their emissions currently are not included in the GHGI. A recent study of 19 abandoned wells in Pennsylvania reported a highly skewed distribution with an average emission rate of 11 g methane h^{-1} (Kang et al. 2014). Due to the large number of abandoned wells, emission rates of this magnitude could result in substantial emissions, but there is high uncertainty because of the small sample size and limited geographic scope of the study. A recent study report emissions from over 100 abandoned wells in four states (Townsend-Small et al. 2016). Geologic microseepage in hydrocarbon-prone basins is estimated to emit 10 Tg methane yr^{-1} globally (Etiope & Klusman 2010). Seepage likely contributes a portion of top-down estimates of fossil methane emissions, but the high variability and lack of data from most regions precludes a reasonable estimate without further research.

Summary

Natural gas has climate benefits compared to more carbon-intensive fossil fuels, but these benefits may be reduced or delayed by methane emissions across the supply chain. U.S. EPA's 2015 GHGI estimate of U.S. natural gas system and petroleum system emissions is 7.3 Tg methane, equivalent to 1.4% of gross gas production. Much of the underlying data of the U.S. EPA GHGI are based on a 1990s study

(Harrison et al. 1996) and therefore may not be representative of technological and operational changes associated with unconventional O&G development. Numerous recent studies have used top-down and bottom-up approaches to quantify O&G methane emissions. Bottom-up studies have estimated that supply chain emissions are roughly in line with GHGI estimates, but with higher or lower emissions from some equipment and sectors. Several top-down studies have reported higher emissions than GHGI estimates, which suggests bottom-up estimates may have remaining uncertainty, particularly from poorly characterized sources, such as storage tanks and gathering pipelines.

U.S. natural gas supply chain methane emissions are likely of a magnitude such that coal-to-gas switching for electricity generation will result in immediate climate benefits, although regional variability in emission rates may cause gas from some basins to be worse for the climate in the short-term. However, methane emissions may be high enough that there is short-term climate damage from other technology conversions, such as diesel-to-gas for heavy-duty vehicles, which require lower supply chain loss rates (Camuzeaux et al. 2015). Regardless of the current emission rate, the climate benefits of natural gas relative to more carbon-intensive fossil fuels can be increased by reducing methane emissions. Fortunately, emissions can be effectively controlled with technologies such as reduced emission completions (Allen et al. 2013). U.S. oil and gas industry methane emissions can be reduced by 40% at a cost of less than $0.01 per thousand cubic feet of produced natural gas (ICF 2014). Several studies have indicated that the majority of emissions come from a small number of sources, many of which are malfunctioning or have otherwise avoidable emissions, but the identity of these sources can be unpredictable. Therefore, comprehensive and frequent leak detection and repair programs to identify and mitigate these sources are critical for reducing emissions.

Figure 1. The short-term climate impacts of natural gas supply chain methane emissions can delay the climate benefits of switching from more carbon-intensive fossil fuels to natural gas. The years until net climate benefits are achieved for three technology conversions (coal-to-gas for electricity, gasoline-to-gas for light-duty vehicles, diesel-to-gas for heavy-duty vehicles) is shown as a function of supply chain natural gas leak rate. For each technology, leak rates at or below the intercept ensure immediate climate benefits (e.g., 0.8% for heavy-duty natural gas vehicles).
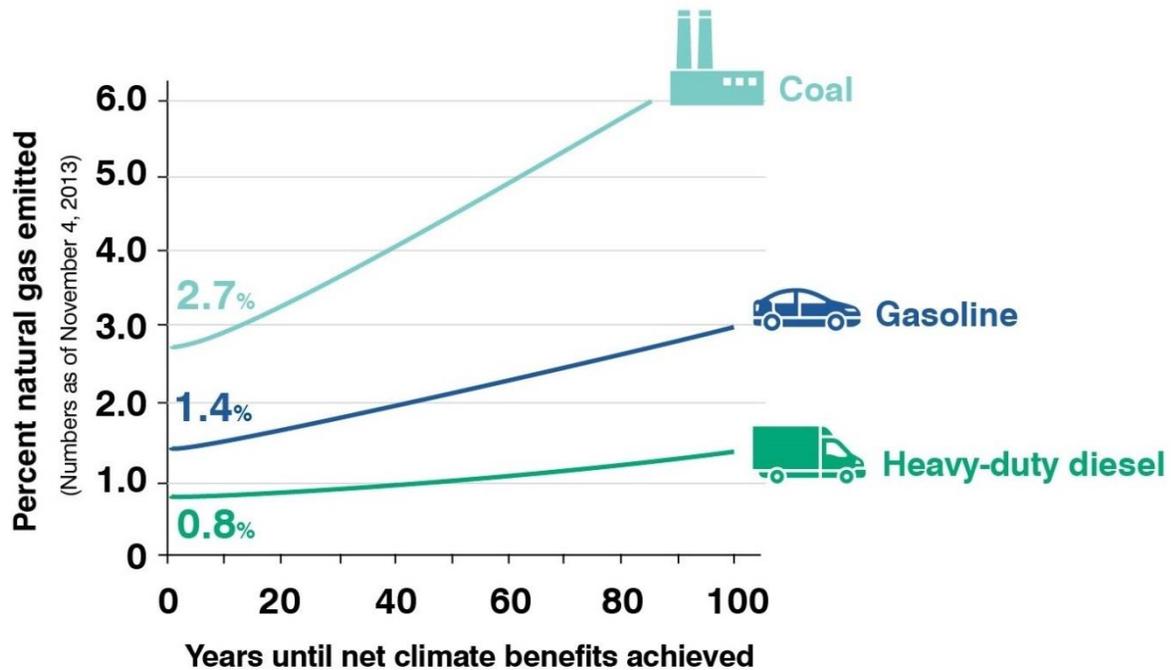
*Adapted from Alvarez et al. (2012) PNAS, **109**: 6435–6440, reflecting new IPCC AR5 & 2013 EPA GHG data. IPCC updates: (1) direct/indirect radiative forcing of CH_4 and CO_2, (2) CH_4 lifetime, (3) CO_2 impulse response function. Additional effects due to climate-carbon feedbacks and CO_2 from the oxidation of CH_4 not included (AR5 lacks data to support time-dependent analysis but EDF believes these effects to be small). Emissions updates include factors in Table 1 and corresponding L_{REF} values in Table S1 of PNAS paper; an L_{REF} value specific to heavy-duty CNG vehicles is now used.*

Figure 2. Emission rate distributions from four source types are plotted as the percent of sites in ascending order of emission rate versus the percent of total emissions from sites at or below that rank. For example, the lowest emitting 50% of well pads contribute 1% of total emissions from measured sites, while the highest emitting 10% contribute 69% of total emissions. Data are from four recent studies (Rella et al. 2015; Subramanian et al. 2015; Mitchell et al. 2015; Lamb et al. 2015).
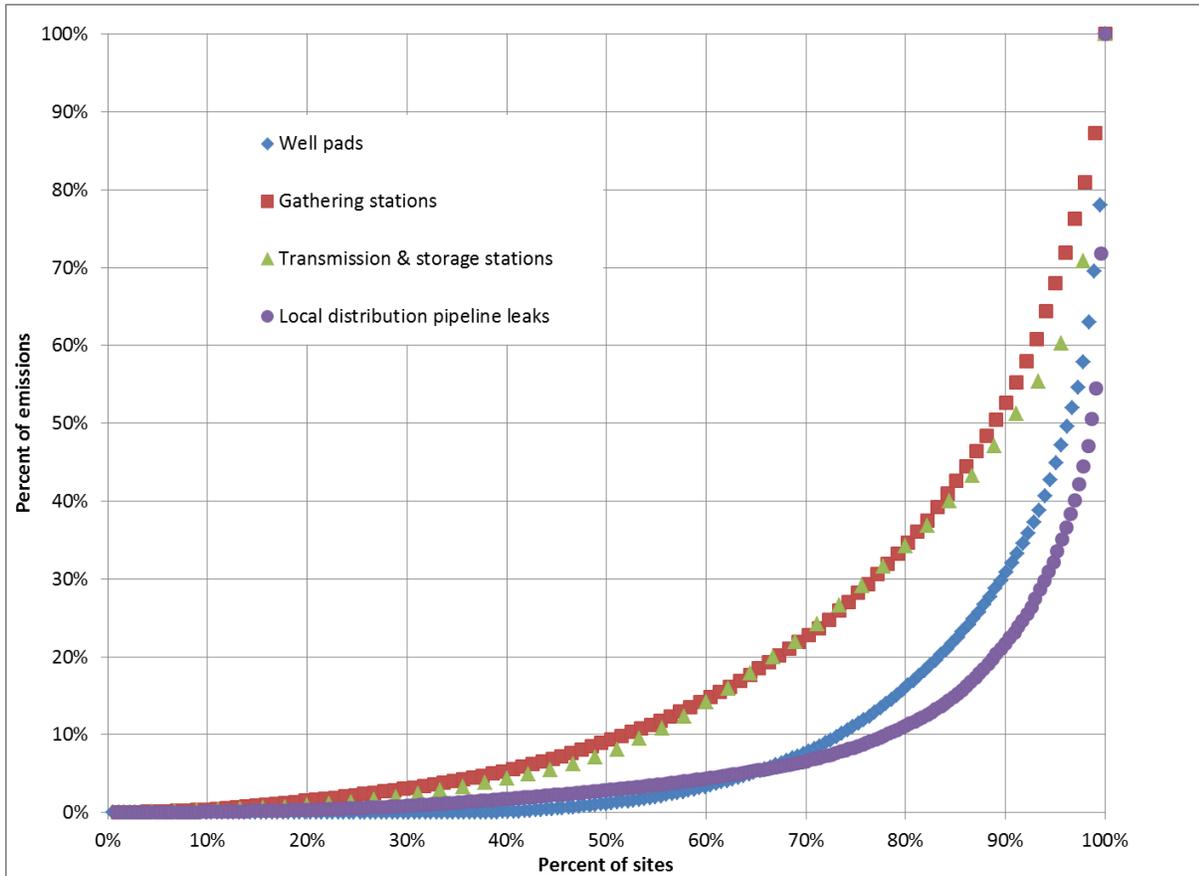
CHAPTER 1 REFERENCES

Allen, D., Torres, V., Thomas, J., Sullivan, D., Harrison, M., Hendler, A., Herndon, S., Kolb, C., Fraser, M., Hill, A., Lamb, B., Miskimins, J., Sawyer, R. and Seinfeld, J. (2013). Measurements of methane emissions at natural gas production sites in the United States. Proceedings of the National Academy of Sciences, 110(44), pp.17768-17773.

Allen, D., Pacsi, A., Sullivan, D., Zavala-Araiza, D., Harrison, M., Keen, K., Fraser, M., Daniel Hill, A., Sawyer, R. and Seinfeld, J. (2015). Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Pneumatic Controllers. Environmental Science & Technology, 49(1), pp.633-640.

Allen, D., Sullivan, D., Zavala-Araiza, D., Pacsi, A., Harrison, M., Keen, K., Fraser, M., Daniel Hill, A., Lamb, B., Sawyer, R. and Seinfeld, J. (2015). Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Liquid Unloadings. Environmental Science & Technology, 49(1), pp.641-648.

Alvarez, R., Pacala, S., Winebrake, J., Chameides, W. and Hamburg, S. (2012). Greater focus needed on methane leakage from natural gas infrastructure. Proceedings of the National Academy of Sciences, 109(17), pp.6435-6440.

Brandt, A., Heath, G., Kort, E., O'Sullivan, F., Petron, G., Jordaan, S., Tans, P., Wilcox, J., Gopstein, A., Arent, D., Wofsy, S., Brown, N., Bradley, R., Stucky, G., Eardley, D. and Harriss, R. (2014). Methane Leaks from North American Natural Gas Systems. Science, 343(6172), pp.733-735.

Brantley, H., Thoma, E., Squier, W., Guven, B. and Lyon, D. (2014). Assessment of Methane Emissions from Oil and Gas Production Pads using Mobile Measurements. Environmental Science & Technology, 48(24), pp.14508-14515.

Camuzeaux, J., Alvarez, R., Brooks, S., Browne, J. and Sterner, T. (2015). Influence of Methane Emissions and Vehicle Efficiency on the Climate Implications of Heavy-Duty Natural Gas Trucks. Environmental Science & Technology, 49(11), pp.6402-6410.

Caulton, D., Shepson, P., Cambaliza, M., McCabe, D., Baum, E. and Stirm, B. (2014). Methane Destruction Efficiency of Natural Gas Flares Associated with Shale Formation Wells. Environmental Science & Technology, 48(16), pp.9548-9554.

Clark, N. N., McKain, D. L., Wayne, W. S., Li, H., Johnson, D. R., Akkerman, V., Sandoval, S., Covington, A. N., Mongold, R. A., Ugarte, O. J. (*In review*). Development of a Methane Emissions Estimation Model for the Heavy-Duty Transportation Sector.

Eastern Research Group. (2011). *City of Fort Worth Natural Gas Air Quality Study*; City of Fort Worth: Fort Worth, TX; http://fortworthtexas.gov/gaswells/air-quality-study/final/

Etiope, G. and Klusman, R. (2002). Geologic emissions of methane to the atmosphere. Chemosphere, 49(8), pp.777-789.

European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL). (2011). Emission Database for Global Atmospheric Research, release version 4.2; http://edgar.jrc.ec.europa.eu

Harrison, M. R.; Shires, T.; Wessels, J. K.; Cowgill, R. M. (1996). Methane emissions from the natural gas industry; EPA: Washington, DC; http://www.epa.gov/methane/gasstar/documents/emissions_report/1_executiveummary.pdf

Harriss, R., Alvarez, R., Lyon, D., Zavala-Araiza, D., Nelson, D. and Hamburg, S. (2015). Using Multi-Scale Measurements to Improve Methane Emission Estimates from Oil and Gas Operations in the Barnett Shale Region, Texas. Environmental Science & Technology, 49(13), pp.7524-7526.

Howarth, R., Santoro, R. and Ingraffea, A. (2011). Methane and the greenhouse-gas footprint of natural gas from shale formations. Climatic Change, 106(4), pp.679-690.

ICF. (2014). Economic Analysis of Methane Emission Reduction Opportunities in the U.S. Onshore Oil and Natural Gas Industries. https://www.edf.org/sites/default/files/methane_cost_curve_report.pdf

IPCC. (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Jeong, S., Hsu, Y., Andrews, A., Bianco, L., Vaca, P., Wilczak, J. and Fischer, M. (2013). A multitower measurement network estimate of California's methane emissions. Journal of Geophysical Research: Atmospheres, 118(19), pp.11,339-11,351.

Johnson, D., Covington, A. and Clark, N. (2015). Methane Emissions from Leak and Loss Audits of Natural Gas Compressor Stations and Storage Facilities. Environmental Science & Technology, 49(13), pp.8132-8138.

Karion, A., Sweeney, C., Pétron, G., Frost, G., Michael Hardesty, R., Kofler, J., Miller, B., Newberger, T., Wolter, S., Banta, R., Brewer, A., Dlugokencky, E., Lang, P., Montzka, S., Schnell, R., Tans, P., Trainer, M., Zamora, R. and Conley, S. (2013). Methane emissions estimate from airborne measurements over a western United States natural gas field. Geophys. Res. Lett., 40(16), pp.4393-4397.

Karion, A., Sweeney, C., Kort, E., Shepson, P., Brewer, A., Cambaliza, M., Conley, S., Davis, K., Deng, A., Hardesty, M., Herndon, S., Lauvaux, T., Lavoie, T., Lyon, D., Newberger, T., Pétron, G., Rella, C., Smith, M., Wolter, S., Yacovitch, T. and Tans, P. (2015). Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region. Environmental Science & Technology, 49(13), pp.8124-8131.

Kang, M., Kanno, C., Reid, M., Zhang, X., Mauzerall, D., Celia, M., Chen, Y. and Onstott, T. (2014). Direct measurements of methane emissions from abandoned oil and gas wells in Pennsylvania. Proceedings of the National Academy of Sciences, 111(51), pp.18173-18177.

Kort, E., Frankenberg, C., Costigan, K., Lindenmaier, R., Dubey, M. and Wunch, D. (2014). Four corners: The largest US methane anomaly viewed from space. Geophys. Res. Lett., 41(19), pp.6898-6903.

Lamb, B., Edburg, S., Ferrara, T., Howard, T., Harrison, M., Kolb, C., Townsend-Small, A., Dyck, W., Possolo, A. and Whetstone, J. (2015). Direct Measurements Show Decreasing Methane Emissions from Natural Gas Local Distribution Systems in the United States. Environmental Science & Technology, 49(8), pp.5161-5169.

Lan, X., Talbot, R., Laine, P. and Torres, A. (2015). Characterizing Fugitive Methane Emissions in the Barnett Shale Area Using a Mobile Laboratory. Environmental Science & Technology, 49(13), pp.8139-8146.

Lavoie, T., Shepson, P., Cambaliza, M., Stirm, B., Karion, A., Sweeney, C., Yacovitch, T., Herndon, S., Lan, X. and Lyon, D. (2015). Aircraft-Based Measurements of Point Source Methane Emissions in the Barnett Shale Basin. Environmental Science & Technology, 49(13), pp.7904-7913.

Lyon, D., Zavala-Araiza, D., Alvarez, R., Harriss, R., Palacios, V., Lan, X., Talbot, R., Lavoie, T., Shepson, P., Yacovitch, T., Herndon, S., Marchese, A., Zimmerle, D., Robinson, A. and Hamburg, S. (2015). Constructing a Spatially Resolved Methane Emission Inventory for the Barnett Shale Region. Environmental Science & Technology, 49(13), pp.8147-8157.

McKain, K., Down, A., Raciti, S., Budney, J., Hutyra, L., Floerchinger, C., Herndon, S., Nehrkorn, T., Zahniser, M., Jackson, R., Phillips, N. and Wofsy, S. (2015). Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. Proceedings of the National Academy of Sciences, 112(7), pp.1941-1946.

Marchese, A. J., Vaughn, T. L., Zimmerle, D., Martinez, D. M., Williams, L. L., Robinson, A. L., Mitchell, A. L., Subramanian, R., Tkacik, D. S., Roscioli, J. R., & Herndon, S. C. (2015). Methane emissions from United States natural gas gathering and processing. Environmental Science & Technology, 49(17), pp.10718-10727.

Miller, S., Wofsy, S., Michalak, A., Kort, E., Andrews, A., Biraud, S., Dlugokencky, E., Eluszkiewicz, J., Fischer, M., Janssens-Maenhout, G., Miller, B., Miller, J., Montzka, S., Nehrkorn, T. and Sweeney, C. (2013). Anthropogenic emissions of methane in the United States. Proceedings of the National Academy of Sciences, 110(50), pp.20018-20022.

Mitchell, A., Tkacik, D., Roscioli, J., Herndon, S., Yacovitch, T., Martinez, D., Vaughn, T., Williams, L., Sullivan, M., Floerchinger, C., Omara, M., Subramanian, R., Zimmerle, D., Marchese, A. and Robinson, A. (2015). Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results. Environmental Science & Technology, 49(5), pp.3219-3227.

Nathan, B., Golston, L., O'Brien, A., Ross, K., Harrison, W., Tao, L., Lary, D., Johnson, D., Covington, A., Clark, N. and Zondlo, M. (2015). Near-Field Characterization of Methane Emission Variability from a Compressor Station Using a Model Aircraft. Environmental Science & Technology, 49(13), pp.7896-7903.

Peischl, J., Ryerson, T., Brioude, J., Aikin, K., Andrews, A., Atlas, E., Blake, D., Daube, B., de Gouw, J., Dlugokencky, E., Frost, G., Gentner, D., Gilman, J., Goldstein, A., Harley, R., Holloway, J., Kofler, J., Kuster, W., Lang, P., Novelli, P., Santoni, G., Trainer, M., Wofsy, S. and Parrish, D. (2013). Quantifying sources of methane using light alkanes in the Los Angeles basin, California. Journal of Geophysical Research: Atmospheres, 118(10), pp.4974-4990.

Pétron, G., Frost, G., Miller, B., Hirsch, A., Montzka, S., Karion, A., Trainer, M., Sweeney, C., Andrews, A., Miller, L., Kofler, J., Bar-Ilan, A., Dlugokencky, E., Patrick, L., Moore, C., Ryerson, T., Siso, C., Kolodzey, W., Lang, P., Conway, T., Novelli, P., Masarie, K., Hall, B., Guenther, D., Kitzis, D., Miller, J., Welsh, D., Wolfe, D., Neff, W. and Tans, P. (2012). Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study. Journal of Geophysical Research: Atmospheres, 117(D4).

Pétron, G., Karion, A., Sweeney, C., Miller, B., Montzka, S., Frost, G., Trainer, M., Tans, P., Andrews, A., Kofler, J., Helmig, D., Guenther, D., Dlugokencky, E., Lang, P., Newberger, T., Wolter, S., Hall, B., Novelli, P., Brewer, A., Conley, S., Hardesty, M., Banta, R., White, A., Noone, D., Wolfe, D. and Schnell, R. (2014). A new look at methane and nonmethane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin. Journal of Geophysical Research: Atmospheres, 119(11), pp.6836-6852.

United States Government Publishing Office. (2015). Electronic Code of Federal Regulations Title 40 Part 98 - Mandatory Greenhouse Gas Reporting. http://www.ecfr.gov/cgi-bin/text-idx?tpl=/ecfrbrowse/Title40/40cfr98_main_02.tpl

United States Government Publishing Office. (2015). Electronic Code of Federal Regulations Title 40 Part 60 Subpart OOOO .Standards of Performance for New Stationary Sources: Subpart OOOO—Standards of Performance for Crude Oil and Natural Gas Production, Transmission and Distribution. http://www.ecfr.gov/cgi-bin/text-idx?SID=true&node=sp40.7.60.oooo

Rella, C., Tsai, T., Botkin, C., Crosson, E. and Steele, D. (2015). Measuring Emissions from Oil and Natural Gas Well Pads Using the Mobile Flux Plane Technique. Environmental Science & Technology, 49(7), pp.4742-4748.

Roscioli, J., Yacovitch, T., Floerchinger, C., Mitchell, A., Tkacik, D., Subramanian, R., Martinez, D., Vaughn, T., Williams, L., Zimmerle, D., Robinson, A., Herndon, S. and Marchese, A. (2015). Measurements of methane emissions from natural gas gathering facilities and processing plants: measurement methods. Atmospheric Measurement Techniques, 8(5), pp.2017-2035.

Shoemaker, J., Schrag, D., Molina, M. and Ramanathan, V. (2013). What Role for Short-Lived Climate Pollutants in Mitigation Policy?. Science, 342(6164), pp.1323-1324.

Simpson, D. (2014). Pneumatic Controllers in Upstream Oil and Gas. Oil and Gas Facilities, 3(05), pp.083-096.

Smith, M., Kort, E., Karion, A., Sweeney, C., Herndon, S. and Yacovitch, T. (2015). Airborne Ethane Observations in the Barnett Shale: Quantification of Ethane Flux and Attribution of Methane Emissions. Environmental Science & Technology, 49(13), pp.8158-8166.

Subramanian, R., Williams, L., Vaughn, T., Zimmerle, D., Roscioli, J., Herndon, S., Yacovitch, T., Floerchinger, C., Tkacik, D., Mitchell, A., Sullivan, M., Dallmann, T. and Robinson, A. (2015). Methane Emissions from Natural Gas Compressor Stations in the Transmission and Storage Sector: Measurements and Comparisons with the EPA Greenhouse Gas Reporting Program Protocol. Environmental Science & Technology, 49(5), pp.3252-3261.

Townsend-Small, A., Tyler, S., Pataki, D., Xu, X. and Christensen, L. (2012). Isotopic measurements of atmospheric methane in Los Angeles, California, USA: Influence of "fugitive" fossil fuel emissions. Journal of Geophysical Research: Atmospheres, 117 D07308, 1-11.

Townsend-Small, A., Marrero, J., Lyon, D., Simpson, I., Meinardi, S. and Blake, D. (2015). Integrating Source Apportionment Tracers into a Bottom-up Inventory of Methane Emissions in the Barnett Shale Hydraulic Fracturing Region. Environmental Science & Technology, 49(13), pp.8175-8182.

Townsend-Small, A., Ferrara, T.W., Lyon, D.R., Fries, A.E. and Lamb, B.K., 2016. Emissions of coalbed and natural gas methane from abandoned oil and gas wells in the United States. Geophysical Research Letters, 43(5) pp.2283-2290.

United States Energy Information Administration. (2015). Washington, DC. http://www.eia.gov/

United States Environmental Protection Agency. (2015). Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2013; EPA: Washington, DC. http://www.epa.gov/ climatechange/ ghgemissions/ usinventoryreport.html

United States Environmental Protection Agency. (2015). Greenhouse Gas Reporting Program; EPA: Washington, DC. http://ghgdata.epa.gov/ghgp/main.do

United States Environmental Protection Agency. (1995) AP-42 Compilation of Air Pollutant Emission Factors ; EPA: Washington, DC; http://www.epa.gov/ttnchie1/ap42/

Wecht, K., Jacob, D., Frankenberg, C., Jiang, Z. and Blake, D. (2014). Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data. Journal of Geophysical Research: Atmospheres, 119(12), pp.7741-7756.

Yacovitch, T., Herndon, S., Pétron, G., Kofler, J., Lyon, D., Zahniser, M. and Kolb, C. (2015). Mobile Laboratory Observations of Methane Emissions in the Barnett Shale Region. Environmental Science & Technology, 49(13), pp.7889-7895.

Zavala-Araiza, D., Lyon, D., Alvarez, R., Palacios, V., Harriss, R., Lan, X., Talbot, R. and Hamburg, S. (2015). Toward a Functional Definition of Methane Super-Emitters: Application to Natural Gas Production Sites. Environmental Science & Technology, 49(13), pp.8167-8174.

Zimmerle, D., Williams, L., Vaughn, T., Quinn, C., Subramanian, R., Duggan, G., Willson, B., Opsomer, J., Marchese, A., Martinez, D. and Robinson, A. (2015). Methane Emissions from the Natural Gas Transmission and Storage System in the United States. Environmental Science & Technology, DOI: 10.1021/acs.est.5b01669.

CHAPTER 2

Constructing a spatially resolved methane emission inventory for the Barnett Shale region

Abstract

Methane emissions from the oil and gas industry (O&G) and other sources in the Barnett Shale region were estimated by constructing a spatially resolved emission inventory. Eighteen source categories were estimated using multiple datasets, including new empirical measurements at regional O&G sites and a national study of gathering and processing facilities. Spatially-referenced activity data were compiled from federal and state databases and combined with O&G facility emission factors calculated using Monte Carlo simulations that account for high emission sites representing the very upper portion, or fat-tail, in the observed emissions distributions. Total methane emissions in the 25-county Barnett Shale region in October 2013 were estimated to be 72,300 (63,400—82,400) kg CH_4 hr^{-1}. O&G emissions were estimated to be 46,200 (40,000—54,100) kg CH_4 hr^{-1} with 19% of emissions from fat-tail sites representing less than 2% of sites. Our estimate of O&G emissions in the Barnett Shale region was higher than alternative inventories based on the United States Environmental Protection Agency (EPA) Greenhouse Gas Inventory, EPA Greenhouse Gas Reporting Program, and Emissions Database for Global Atmospheric Research by factors of 1.5, 2.7, and 4.3, respectively. Gathering compressor stations, which accounted for 40% of O&G emissions in our inventory, had the largest difference from emission estimates based on EPA data sources. Our inventory's higher O&G emission estimate was due primarily to its more comprehensive activity factors and inclusion of emissions from fat-tail sites.

Introduction

Fossil fuel substitutions resulting from the recent growth of natural gas production have the potential to immediately reduce CO_2 emissions and long-term climate impacts, but emissions of methane from the natural gas supply chain may also increase short-term climate impacts.[1] Several recent studies have used different methodologies to estimate the magnitude of oil and gas industry (O&G) methane emissions.[2–6] Top-down approaches, which quantify emissions from a region using atmospheric measurements of well mixed air, have inferred higher O&G methane emissions than bottom-up approaches, which estimate regional emissions by constructing inventories based on activity factors and emission factors.[7,8] Reported differences may result in part from top-down studies incorrectly attributing emissions to O&G sources or sampling during times when short-term events are occurring at a different rate than predicted by inventories. Additionally, bottom-up studies may underestimate emissions due to incomplete activity factors or emission factors based on measurements that exclude the fat-tail of a skewed emission rate distribution — relatively rare sources that contribute a large fraction of total emissions. Coordinated top-down and bottom-up measurements are needed to reconcile the two methods and more accurately estimate methane emissions.[8,9] Development of a detailed emission inventory composed of both more complete activity factors and more representative emission factors is a critical step in top-down/bottom-up reconciliation.

The Barnett Shale of north-central Texas was the first shale basin to be developed for natural gas with a combination of horizontal drilling and hydraulic fracturing. Development peaked in 2008 with over 4,000 drilling permits issued, then declined to less than 1,000 issued permits 2013.[10] Barnett Shale production peaked at 5.7 billion cubic feet (Bcf) natural gas day^{-1} in 2012 and 28 thousand barrels (Mbbl) day^{-1} hydrocarbon liquids (oil and natural gas condensate) in 2011[10], but the basin is expected to remain a major contributor to U.S. natural gas production through 2030[11]. A mature field is ideal for investigating long-term methane emissions from O&G sites in the production phase but provides fewer opportunities to observe emission events from well development activities, which can be challenging to characterize due to their short duration and spatiotemporal heterogeneity[12].

The Barnett Shale region contains most of the Dallas-Fort Worth-Arlington Metropolitan Statistical Area, which has a population over 6.5 million and includes many urban methane sources such as landfills. The region also includes extensive rural land use and over one million cattle. O&G air pollution sources in the region have been extensively studied; for example, a criteria and hazardous air pollutant emission inventory was developed by the Texas Commission on Environmental Quality (TCEQ)[13] and field measurements of pollutant emissions, including methane, were commissioned by the City of Fort Worth[14].

During the two week period of October 16—30, 2013, ten research teams performed multi-scale measurements in the Barnett Shale region to quantify methane emissions from O&G and other sources including landfills (the Barnett Coordinated Campaign). This paper uses bottom-up measurements from the Barnett Coordinated Campaign and other available data to construct a spatially resolved methane emission inventory (4 km x 4 km grid cells) for the 25-county Barnett Shale region defined by the Texas Railroad Commission[10]. Natural gas production site emission estimates were characterized in Zavala-Araiza et al.[15] Our bottom-up emission estimates were compared to alternative emission inventories we developed from commonly cited sources: the United States Environmental Protection Agency (EPA) Greenhouse Gas Reporting Program (GHGRP)[16], EPA United States Greenhouse Gas Inventory (GHGI)[17], and the Emissions Database for Global Atmospheric Research v4.2 (EDGAR)[18]. These inventories have been shown to produce lower methane emissions rates than top-down studies regionally and nationally[4,7,8,19]. The top-down estimates made as part of the Barnett Coordinated Campaign and an additional two week period in March 2013 included the core production area of the Barnett Shale, but did not include all of the 25-county area of this inventory.[20,21] The gridded inventory constructed for this paper can be used to estimate emissions in other spatial domains in the Barnett region including areas measured by top-down methods.

Methods and Data

A spatially resolved methane emissions inventory for the 25-county Barnett Shale region was constructed using a combination of bottom-up approaches to estimate emissions from O&G and other sources. Emissions from O&G facilities (production sites, compressor stations, and processing plants) were

estimated with emission factors calculated using Monte Carlo simulations, which account for the uncertainty associated with the variability of measured site emission rates[22]. Other emission sources were estimated using data from the GHGRP, GHGI, and published literature. Activity factors were spatially referenced to estimate emissions within grid cells, similar to the approach used in Jeong et al.[23] The 4 km x 4 km grid cells conform to the Comprehensive Air Quality Model Texas domain with extensions.[24] Emissions are grouped into three classes: O&G (active well to customer meter), other thermogenic (fossil sources not included in GHGI natural gas and petroleum systems), and biogenic. Emissions are reported as central estimates with 95th percent confidence intervals; total and category subtotal uncertainties are estimated by quadrature summation of the uncertainties in each source category.

Activity factors

The number and location of O&G and other methane-emitting facilities were compiled from multiple state and federal databases. Facilities with annual greenhouse gas emissions of ≥25,000 metric tons carbon dioxide equivalents (CO_2e) are required to submit annual emissions to the GHGRP.[16] These sites were classified as gas transmission, gas processing, gas gathering, landfills, or other industrial sites based on the GHGRP subpart under which they report. Additional O&G sources were identified using two datasets from the Texas Commission on Environmental Quality (TCEQ): the 2009 Barnett Shale Area Special Inventory (BSASI)[13] and the air permit database[25]. These sites were classified based on equipment type and facility name. Compressor stations were classified as gathering (upstream of processing) or transmission (downstream of processing) based on their proximity to gathering and transmission pipelines. O&G well locations were obtained from DI Desktop[26] and clustered into production sites as described in Zavala-Araiza et al.[15] Google Earth imagery was used to quality control reported spatial coordinates, manually locate sites without reported coordinates, and remove duplicate and decommissioned sites. The Supporting Information (SI) includes additional details on the compilation and classification of activity factors (section SI1), a map (Figure SI1) and spreadsheet with facility locations.

Monte Carlo simulations of O&G emissions

For each O&G facility type, emission factors with a 95th percent confidence interval were calculated with Monte Carlo simulations that drew from facility-specific emission rate distributions assembled from measurements made during the Barnett Coordinated Campaign[27–28] and a recent national study on methane emissions from gathering and processing facilities[31]. Two emission rate distributions were used for each Monte Carlo simulation. The first emission rate distribution, defined as the "sampled distribution", was constructed from data collected by unbiased sampling of the Barnett region or the national population. Due to the positively skewed emission rate distribution of many O&G facility types[3,28,31], the mean emission rate of a random sample may underestimate the average emission rate of the entire population if the sample size is insufficient to fully capture the highest end, or fat-tail, of the distribution. Figure 1 compares the sampled distributions by facility type to emission rates observed during the Barnett Coordinated Campaign using sampling methods biased towards higher emission sites. [27,29,30] These other Barnett datasets include measurements exceeding the maximum of the sampled distributions, which indicates unbiased sampling did not fully capture the fat-tail. To account for the effect of these high-emitting sites, we constructed a second emission rate distribution, defined as the "fat-tail site distribution", from data representing sites with emission rates exceeding the maximum value in the unbiased, sampled distribution of each facility type. A two-step Monte Carlo simulation was performed with the first step drawing from the sampled distribution and the second step drawing from the fat-tail site distribution. For each facility type, the probability of drawing from the fat-tail distribution was a best estimate based on the number of observed fat-tail sites compared to the total sites in the region. We ran sensitivity tests using a range of probabilities (0—5%) to test the effect of this assumption. Each Monte Carlo simulation included 10,000 iterations of random selection with replacement from one of the two emission rate distributions for every facility in the 25-county region. The 95th percent confidence interval of regional emission estimates was determined by the 2.5[th] percentile and 97.5[th] percentile of the 10,000 iterations. Facility-specific emission factors were calculated for each facility type by dividing the regional emission estimates by the number of facilities in the region (Figure SI2 illustrates the method). Spatially resolved O&G facility emissions were estimated by applying the emission factors to the spatially-referenced activity data.

Compressor station emissions were estimated with two-phase Monte Carlo simulations drawing from site emission rate distributions constructed using data from a national study of gathering and processing

facilities(Mitchell et al.)[30] and the Barnett Coordinated Campaign[27,29,30]. Gathering stations comprised over 90% of the compressor stations in the region. Transmission stations and storage facilities were treated identically to gathering compressor stations since they have similar equipment and installed engine horsepower. The sampled distribution was constructed using a national dataset of 100 gathering stations with compression or a combination of compression and dehydration equipment with site emissions ranging from of 0 to 700 kg CH_4 hr^{-1} and averaging 55 kg CH_4 hr^{-1}.[30] The fat-tail site distribution included four gathering sites measured during the campaign ranging from 1,360 to 2,120 kg CH_4 hr^{-1}.[27,29,30] The probability to draw from the fat-tail site distribution was set at 1%, which is equivalent to 2 to 3 compressor stations in the Barnett region with fat-tail emission rates at any moment in time. This probability was chosen based on the observation of four fat-tail sites over a 15-day period out of a population of 276 facilities. Multiple simulations were run with the probability of selecting from the fat-tail site distribution ranging from 0 to 5% to test the sensitivity of the outcome to the 1% assumption.

Processing plant emissions were estimated following a similar approach as compressor stations. Monte Carlo simulations drew from two sets of site emission rate distributions constructed using data from Mitchell et al.[31] for the unbiased sample distribution and the Barnett Coordinated Campaign[29,30] for the fat-tail site distribution. Because of the wide variation in processing plant size and complexity, processing plants were subdivided into two classes: large if they reported to the GHGRP or small if they did not. The average installed horsepower of Barnett plants (large = 21,000 HP, small = 8,000 HP) supports this division. The sampled distribution for large processing plants was constructed from a national dataset of 16 processing plants with site emissions ranging from 4 to 600 kg CH_4 hr^{-1} and averaging 170 kg CH_4 hr^{-1}.[31] The sampled distribution for small processing plants was constructed from a national dataset of nine gathering stations containing a combination of compression, dehydration, and treatment (C/D/T) equipment with site emissions ranging from 7 to 240 kg CH_4 hr^{-1} and averaging 78 kg CH_4 hr^{-1}.[31] Although C/D/T sites were not defined as processing plants by Mitchell et al., they have similarities to small plants including gas treatment and comparable installed horsepower (5000 HP).[31] Five processing plants measured during the campaign were used in the fat-tail distributions.[29,30] For large plants, the fat-tail distribution used two measurements exceeding the sampled distribution (750 and 1,720 kg CH_4 hr^{-1}). For

small plants, the fat-tail distribution used three measurements exceeding the sampled distribution (320, 390, and 490 kg CH_4 hr^{-1}). The two higher values in the large plant fat-tail distribution were not used for small plants because they would require unreasonably high leak rates for these smaller throughput facilities. The probability of selecting from the fat-tail site distributions was set at 2%, which is equivalent to a single processing plant in the region with a fat-tail emission rate at any one moment. Multiple simulations were run with the probability of selecting from the fat-tail site distributions ranging from 0 to 5% to test the sensitivity of the outcome to the 2% assumption.

Production site emissions were estimated using a more complex approach that defined fat-tail sites based on proportional loss rates (methane emitted relative to methane produced). The method is briefly described below with additional details in the SI (section S2); the full method and results are found in Zavala-Araiza et al.[15] Activity factors were based on estimated O&G production site counts. Emission factors were derived with Monte Carlo simulations drawing from site emission rate distributions constructed using data from 226 sites measured during the Barnett Coordinated Campaign.[27–29] Activity and measurement data were divided into cohorts based on gas production and production-normalized emissions. As described in Zavala-Araiza et al.[15], the sites with the highest proportional loss rates were defined as γ-sites; a fat-tail probability of 0.25% for γ-sites was chosen and a sensitivity analysis was performed to test the effect of differing probabilities on estimated emissions. Zavala-Araiza et al.[15] reports emissions only for gas-producing sites. For this paper, gas-producing site emissions were divided into gas sites and oil sites based on the well type reported in DI Desktop[26]. In addition, emissions from oil sites with no gas production were estimated using an emission factor of 5.14 x 10^{-3} kg CH_4 hr^{-1} $well^{-1}$ based on the Petroleum Systems stripper well emission factor in the GHGI.[17]

Other O&G sources
Production site emissions estimated with the Monte Carlo simulations only included emissions during the operation phase. Additional emissions can occur episodically during drilling, completion flowback, or maintenance activities. Completion flowback emissions, which occur when a well is vented after hydraulic fracturing to prepare for routine production, were estimated for 73 individual well completions that

occurred during the Barnett Coordinated Campaign based on well locations and completion start dates from DI Desktop[26]. In summary, emissions were estimated based on initial gas production with an assumption that gas wells, but not oil wells, controlled emissions due to federal regulations[32]. The average of the daily completion emission estimates during the campaign was used as the central estimate and the minimum and maximum daily estimates were used as the lower and upper bound estimates. The detailed methods are described in SI Section SI3.

Gathering and transmission pipeline emissions were estimated from pipeline mileage and per mile emission factors. GIS shapefiles of gathering and transmission pipelines from DI Desktop[26] were joined with the grid to determine the miles of transmission and gathering pipelines in each grid cell. Emissions were estimated using the GHGI emission factors from the production and transmission & storage sectors.[17] Gathering pipelines used an emission factor of 4.7×10^{-2} kg CH_4 hr^{-1} $mile^{-1}$ based on Midcontinent production sector emission factors for pipeline leaks, pipeline blowdowns, and mishaps. Transmission pipelines used an emission factor of 7.1×10^{-2} kg CH_4 hr^{-1} $mile^{-1}$ based on transmission and storage sector emission factors for pipeline leaks and pipeline venting. Uncertainty was based on EPA's uncertainty estimate (+30%/-19%) for GHGI Natural Gas Systems.[17]

Natural gas distribution emissions were estimated using data from a recent national study of methane emissions from local distribution pipelines and metering and regulating (M&R) stations (Lamb et al.). [33] In summary, activity factors were based on data reported by Atmos Energy, which is the utility serving the vast majority of customers in the Barnett region. Emissions from sources not measured in Lamb et al.[33] were estimated using GHGI national emissions[17] prorated by activity factors. The detailed methods are described in SI Section SI4. The upper confidence limit uncertainty (+71%) was based on the emission factor uncertainties of Lamb et al.[33]; for the lower confidence limit, EPA's uncertainty estimate (-19%) for GHGI Natural Gas Systems[17] was used since Lamb et al. only reports upper confidence limits.

Other thermogenic sources

Abandoned well emissions were estimated using well counts and a per well emission factor. The locations of inactive and plugged and abandoned wells in the Barnett region were obtained from DI Desktop.[26] For the subset of wells without coordinates, activity data were aggregated by county. The emission factor and uncertainty is based on the observed average emission rate of nine abandoned wells in the Marcellus Shale, 1.1×10^{-2} (+100%/-50%) kg CH_4 hr^{-1} $well^{-1}$.[34]

Emissions from other industrial sources reporting to the GHGRP were based on reported 2013 emissions.[16] Annual emissions were converted to kg CH_4 hr^{-1} by assuming a constant emission rate. Uncertainty for industrial source emissions was assumed to be +138%/-58%, which is the uncertainty of the combustion emission factor used to estimate GHGRP emissions.[35]

Residential and commercial end use emissions from leaks past the meter and incomplete combustion of natural gas by heaters and appliances were estimated using October 2013 gas deliveries to residential and commercial customers. Barnett gas consumption was estimated by prorating statewide monthly gas deliveries[36] by 2010 population[37]. For the central estimate, it was assumed that 0.16% of delivered gas was emitted, which is based on measurements of five California residences.[38] For the lower bound, a leak rate of 0.028% was based on the GHGI emission factor for residential and commercial stationary combustion.[17] For the upper bound, a leak rate of 1.6% was based on a Boston study that reported 2.7% of delivered gas was emitted – a state emission inventory estimated emissions from other sources in the Boston region were equivalent to 1.1% of delivered gas, so this leak rate assumes that the remainder of emissions in that study were due to residential and commercial end use.[39]

Methane emissions from gasoline and diesel onroad vehicle were based on county-level annual emissions reported in the 2011 National Emissions Inventory.[40] These emissions were estimated by the EPA using the MOVES2010b model.[41] No data were found on the uncertainty of these emission estimates so uncertainty was conservatively excluded.

Natural gas vehicle emissions were estimated at the county-level based on the volume of natural gas delivered as vehicle fuel in October 2013 and assumed leak rates. The state-wide fuel delivery (210 MMscf)[36] was prorated by county-level vehicle miles traveled.[42] Emissions were assumed to equal 1% of fuel delivered with an uncertainty bound of 0.5—5%. This assumption is highly uncertain but has minor impact on the overall inventory due to the low usage of natural gas vehicles.

Geologic seepage emissions were estimated using a per area emission factor of 0.0184 kg CH_4 hr^{-1} km^{-2}, which is based on a global average net flux of 4.42 mg CH_4 day^{-1} m^{-2} for microseepage and an assumption of 90% methanotrophic consumption.[43] This source category is highly uncertain but no data were found to quantify the uncertainty; therefore, uncertainty was conservatively excluded.

Biogenic sources

Emissions from landfills reporting to the GHGRP were based on reported 2013 emissions.[16] Annual emissions were converted to kg CH_4 hr^{-1} by assuming a constant emission rate. We identified 712 additional landfills by querying TCEQ municipal solid waste permit data.[44] Based on the EPA estimate that 82% of landfill emissions are from facilities reporting to the GHGRP[45], emissions from the non-reporting landfills were estimated by allocating 18% of GHGRP landfill emissions evenly among the 712 facilities. Uncertainty was based on EPA's uncertainty estimate (+49%/-56%) for GHGI landfills.[17] This uncertainty does not account for potential temporal variability in landfill emissions due to factors such as changing atmospheric pressure.

Livestock emissions from cattle manure management and enteric fermentation were estimated using activity data from the United States Department of Agriculture (USDA) and TCEQ and emission factors from the GHGI. Confined animal feeding operation (CAFO) locations and head counts of beef cattle, milking dairy cattle, and non-milking dairy cattle were obtained from the TCEQ water quality general permit database.[46] County-level 2013 head counts of beef cattle, dairy cattle, and unspecified cattle were obtained from the USDA National Agricultural Statistics Service database.[47] Unspecified cattle were assumed to be beef cattle. Dairy and beef cattle populations were further classified into detailed animal

types by assuming the same proportion as the Texas 2012 cattle population used in the GHGI.[17] USDA county-level head counts were adjusted downward to account for the CAFO population in each county, which were treated separately as point sources. Enteric fermentation and manure management emission factors for beef cattle and dairy cattle animal types were derived from the GHGI Texas activity data and emissions.[17] Livestock methane emissions were estimated by multiplying the animal type head counts and the GHGI animal type emission factors for enteric fermentation and manure management. Livestock emission uncertainty was based on the Intergovernmental Panel on Climate Change Tier 2 methodology uncertainty (±20%)[48], which is similar to the approach used in the GHGI[17].

Domestic wastewater treatment emissions were estimated from GHGI 2013 national emissions.[17] Population data from the 2010 US Census were used to prorate national emissions.[37] Population was spatially allocated based on census tract population data.[49] Uncertainty was based on EPA's uncertainty estimate (+2%/-39%) for GHGI Wastewater Treatment.[17]

Other potential methane sources in the region (e.g., reservoirs, wetlands, abandoned coal mines) were assumed to have negligible emissions.

Spatially resolved emission inventory

Emissions data included sources with three levels of spatial resolution. GHGRP facilities, O&G facilities, landfills, and CAFOs were referenced to a specific latitude/longitude. Population-based and area-based emission estimates of natural gas distribution, wastewater treatment, residential and commercial end use, and geologic seepage were attributed to 4 km x 4 km grid cells based on the fractional area and population of each cell. Vehicle and a subset of abandoned well and livestock emissions were estimated at the county level with emissions spatially distributed across the grid proportional to the fraction of county land area in each cell. In addition to generating a gridded emission inventory by source category, emissions were estimated for the 25-county Barnett Shale region based on the spatial intersection of the grid cells and county boundaries.

Alternative emission inventory estimates

Alternative O&G emission inventories were constructed using data from GHGRP, GHGI, and EDGAR. Emission data were scaled to account for the different spatial domains as described briefly below. The GHGRP inventory was based solely on 2013 reported emissions from regional O&G point sources and onshore production basins, which only includes facilities meeting the 25,000 metric ton CO_2e reporting threshold.[16] The GHGI inventory was based on 2013 national emissions from Natural Gas Systems and Petroleum Systems[17] with individual source categories prorated by the ratio of Barnett region and national parameters such as gas production and transmission pipeline miles. The EDGAR inventory was based on EDGAR v4.2 2010 emissions from the gas production/distribution and oil production/refineries sectors.[19] Emissions were converted from 0.1° x 0.1° cells to the 25-county region using the spatial intersection of the cells and county boundaries. EDGAR 2010 emissions were extrapolated to 2013 using the 2013/2010 ratio of Barnett region gas and oil production from DI Desktop[26]. A detailed description of the methods used to construct each methane emission inventory for the 25-county Barnett Shale region in October 2013 is included in SI Section SI5.

Results and Discussion

Barnett region emission estimates

Estimated total emissions in the 25-county Barnett Shale region for October 2013 are 72,300 kg CH_4 hr^{-1} (95th percent confidence interval = 63,400 – 82,400 kg CH_4 hr^{-1}). O&G sources are estimated to emit 46,200 (40,000 – 54,100) kg CH_4 hr^{-1}, or 64% (52 –78%) of total emissions (Table 1). Thermogenic sources, which include additional emissions from abandoned well, natural gas end use, and geologic seepage, are 48,400 (42,100 – 56,400) kg CH_4 hr^{-1}, or 67% (55 – 81%) of total emissions. Gathering compressor stations and active well pads are the largest emission sources, contributing 26% and 25% of total emissions, respectively. Livestock and landfills are the largest biogenic emission sources, contributing about 16% each. A core region of eight counties responsible for 94% of gas production and 43% of oil production contributes 67%, 77%, and 75% of the total, O&G, and thermogenic emissions, respectively, in the Barnett region (Figure 2).

Fat-tail sites contribute 19% (14 – 26%) of O&G emissions and 12% (9 – 15%) of total emissions in our reported inventory estimate, which assumes fat-tail emission rates at 0.25% of production sites[15], 2% of processing plants, and 1% of compressor stations. At these probabilities, there would be approximately 50 production sites, 1 processing plant, and 2 to 3 compressor station with fat-tail emission rates somewhere in the Barnett region at any moment in time. The research teams were able to identify and measure these sites despite their limited numbers in a large region by utilizing specific sampling strategies (e.g., aircraft-based surveys targeting sites with high methane enhancements[30]). O&G site emission factors are dependent on the selected fat-tail site probability (Figure SI3). If the probability of fat-tail sites were reduced by half, O&G emissions would decrease by 8%, while at double the probability, O&G emissions would increase by 16%. Additional sensitivity analyses for production site emissions are reported in Zavala-Araiza et al.[15]

Fat-tail sites do not necessarily have persistently high emissions, but may represent short-term emission events caused by maintenance activities or malfunctions. For production sites, fat-tail γ-sites included emission rates up to 287 kg CH_4 hr^{-1}, approximately six times higher than the maximum emission rate observed using unbiased sampling.[27–29] An effort to identify high emitting sites in the Marcellus Shale region observed average emissions of 850 kg CH_4 hr^{-1} at seven multi-well sites in the drilling phase, which the authors attributed to the conveyance of methane from overlying coal formations through the wellbore.[6] The high emission rates observed during the Barnett Coordinated Campaign do not appear to be related to drilling or hydraulic fracturing due to the infrequent occurrence of these activities during the campaign, but they may be caused by major malfunctions at production sites (e.g., stuck separator dump valve).[15] Another possibility is that measurements occurred during maintenance events such as venting to unload liquids accumulated in the wellbore. The median emission rate of unloading event from 107 wells in a nationwide study was equivalent to 257 kg CH_4 hr^{-1}, similar to our fat-tail production site emission rates.[50] Based on the low number of unloading events reported to the GHGRP in the Barnett region[16], emissions associated with liquids unloading are unlikely to be a major emission source in this case, but may be substantial in regions with frequent unloading events such as the San Juan Basin.

For compressor stations and processing plants, the maximum fat-tail emission rates were 2,040 and 1,720 kg CH_4 hr^{-1}, respectively. These emission rates are higher than the maximum annual average 2013 facility emissions reported to the GHGRP for transmission (520 kg CH_4 hr^{-1}) and processing (1,050 kg CH_4 hr^{-1}).[16] However, the GHGRP reports almost 2,400 unique blowdown (emptying or depressurizing a gas-filled vessel) events nationally exceeding 1,000 kg CH_4 total emissions in 2013, including over 800 events exceeding 10,000 kg CH_4.[16] Since the typical duration of these events range from minutes to hours, short-term blowdown events could cause fat-tail magnitude emission rates observed at compressor stations and processing plants during the Barnett Coordinated Campaign, but the probability of their observation is likely low. Additionally, GHGRP protocols may not capture high emissions from some malfunctions. For example, a recent national study of 45 transmission and storage compressor stations found two sites with emissions up to 1,000 kg hr^{-1} likely caused by leaking isolation valves; the GHGRP-compliant on-site surveys reported emissions 2-3 orders of magnitude lower.[51]

For our study, we define fat-tail sites as those with emission rates above the sampled distribution, but this does not indicate that they are the only high emission sites. The sampled distributions are positively skewed and include sites with high emission rates, some of which had substantial tank venting due to equipment issues.[31] Consequently, there is a larger population of sites than the fat-tail sites in our analysis that contribute a large fraction of regional emissions and have avoidable, excess emissions.[15]

Barnett Shale O&G wells produced 5.6 Bcf day^{-1} natural gas and 54.5 Mbl oil and condensate day^{-1} in October 2013.[26] Assuming a constant production rate and weighted average gas composition of 88.5% methane by volume, our O&G emission estimate is equivalent to 1.2% (1.0 – 1.4%) of gas production. If oil production site emissions (4% of O&G total) are excluded, then the natural gas leak rate decreases to 1.1% (1.0% – 1.3%). Allocating emissions between natural gas and hydrocarbon liquids on an energy basis according to the methods of Zavala-Araiza et al.[52] attributes 95% of emissions to natural gas, resulting in a similar adjusted leak rate.

Comparing inventories

The O&G emission inventory reported here for the Barnett region is a factor of 1.5 (1.3 – 1.7) greater than the emissions estimated from GHGI 2013 national emissions (46,200 versus 31,000 kg CH_4 hr^{-1}; Table 3). This difference is similar to the ratio of ~1.5 between top-down and GHGI estimates of total US methane emissions reported in Miller et al.[7] and Brandt et al.[8] Comparing the inventory reported here to an estimate based on GHGRP 2013 emissions (17,000 kg CH_4 hr^{-1}) yields a factor of 2.7 (2.4 – 3.2) higher emissions from our inventory, comparable to the ratio of 2.5 between aircraft-based mass balance and GHGRP estimates of methane emissions from O&G activities in the Denver-Julesburg basin.[4] Comparing our inventory to an estimate derived from EDGAR 2010 emissions (10,800 kg CH_4 hr^{-1}) yields a factor of 4.3 (3.7 – 5.0) higher emissions from our inventory, similar to the ratio of 4.9±2.6 between O&G emission estimates for the south-central US based on atmospheric data versus EDGAR v4.2.[7]

Production sector emissions based on GHGI and GHGRP emissions are 31% and 21% lower than our estimate, respectively. Comparing the underlying activity factors and emission factors allows for a more detailed assessment of the inventories (Table 3). GHGI and GHGRP site emission factors are not directly reported in these data sources, but derived by dividing total emissions by the number of sites. For the production sector, our activity and emission factors are converted from a per site basis to a per well basis using a factor of 1.4 wells $site^{-1}$ based on the 25-county average. Compared to our estimates for production sites excluding completions, the GHGI activity and emission factors are 11% and 17% lower, respectively. The GHGI emission factor may be lower because the data underlying the GHGI, which is from a 1990s study[53], is not representative of current operational practices. The GHGRP has a 59% lower activity factor and 94% higher emission factor than our estimate. The GHGRP emissions and activity factor are expected to be lower since the data only include facilities meeting the 25,000 metric tons CO_2e reporting threshold. The higher GHGRP emission factor could be due to either reporting facility wells having higher emissions than the regional average or an overestimation of reported emissions caused by GHGRP methods. Our well pad emission factor (0.87 kg hr^{-1}) is between the geometric means of Midcontinent region (0.54 kg hr^{-1}) and Barnett well pads (1.19 kg hr^{-1}) from two recent studies[2,3], supporting the consistency of our estimates.

The largest difference among this work and other inventories is for the gathering sector; the GHGI and GHGRP emissions estimates are a factor of 7.3 and 4,900 lower than our estimate, respectively. Since the GHGI groups gathering within the production sector, we disaggregated emissions by assuming all compressor and pipeline emissions are associated with the gathering sector. The GHGI activity factor for gathering stations, which only includes large stations, greatly underestimates the number of facilities in the Barnett region. An alternative GHGI station activity factor can be estimated from the GHGI production sector compressor engine activity factor by assuming 3.1 compressors per station based on the average from Mitchell et al.[31] This alternative activity factor is three times higher than our facility count, which is probably because the study used to develop the GHGI compressor engine activity factors grouped together production site wellhead compressors and gathering station compressors[53]. Using this high GHGI activity factor, the GHGI emission factor is still 42 times lower than our emission factor. The GHGRP gathering station activity factor and emission factor are 4.8 and over 1,200 times lower than our factors, respectively. GHGRP Subpart W currently exempts gathering stations from reporting vented and fugitive methane emissions.[54] Therefore, GHGRP data only includes gathering facilities reporting combustion emissions under Subpart C .[55] Reporters are required to use a default methane emission factor based on natural gas turbines, which is known to be at least two orders of magnitude too low for reciprocating engines used by the vast majority of gathering stations.[56]

For the processing sector, GHGI emissions are 58% higher than our estimate. The GHGI emission factor is within 10% of our average processing plant emission factor; therefore, the higher emissions are caused primarily by a 45% higher activity factor. GHGRP emissions are a factor of 6.9 lower than our estimate – some of this difference is due to 16 of 38 plants that are not required to report to GHGRP, but the largest difference is from the emission factors. The average reported emissions of GHGRP processing plants is a factor of 5.5 lower than our large processing plant emission factor, which may be due to the exclusion of certain emission sources (e.g., tanks) from GHGRP reporting requirements.[54]

For the transmission and storage sector, the inventory comparisons have similar trends as processing. GHGI emissions are 50% higher than our estimate. The GHGI transmission and storage compressor

station activity factor and emission factor (excluding pipeline emissions) are 47% and 43% higher than our factors, respectively. GHGRP emissions are a factor of 6.0 lower than our estimate, which primarily is due to about 70% of facilities not being required to report to the GHGRP. The average reported emissions of GHGRP facilities is 15% lower than our emission factor, which may be due to the exclusion of certain compressor engine operating mode emissions from GHGRP reporting requirements[54]. Our compressor station emission factor is within 10% of the value used for California compressor stations in Jeong et al.[23] The average emissions of five compressor stations and storage facilities measured during the Barnett Coordinated Campaign with on-site, leak and loss audits was 19 kg CH_4 hr^{-1} .[57] This value is a factor of 3.8 lower than our emission factor, but within 30% of the median value of our compressor station sampled distribution, which demonstrates how emissions could be underestimated if an emission factor is based on a small sample size of a skewed distribution.

For the distribution sector, our emission estimates are a factor of 4.7 and 1.5 lower than the GHGI and GHGRP estimates, respectively. This is due to our study using emission factors from a recent national study that reported emissions from pipelines and M&R stations have decreased since the 1990s when a previous study collected measurements used to develop the GHGI and GHGRP emission factors.[53]

Our detailed, spatially explicit methane emission inventory for the Barnett Shale region illustrates the limitations of relying on commonly used data sources such as GHGI and GHGRP to estimate regional emissions. The GHGI Natural Gas Systems section relies primarily on national emission factors developed in the 1990s to estimate natural gas industry emissions[53] and may not reflect regional differences or recent changes in emission profiles. The GHGRP only includes emissions from facilities meeting a reporting threshold and excludes most emissions from the gathering sector and certain emission sources; therefore, it is inherently an underestimate of emissions and should not be viewed as a complete emission inventory. EPA has recently made changes to improve the completeness of the GHGI and GHGRP and has proposed adding gathering facilities to the GHGRP[58].

Our methane inventory estimates higher emissions than other inventories predominantly due to two reasons: more complete, region-specific activity factors and the inclusion of emissions from fat-tail sites. Our comprehensive search of multiple data sources revealed a substantially higher count of O&G facilities than was contained in any single data source, particularly with regards to gathering compressor stations. Relatively rare, high emitting fat-tail sites such as those observed during the Barnett Coordinated Campaign were estimated to contribute 19% of regional O&G emissions. Our estimate of total methane emissions from the 25-county Barnett Shale region, 72,300 (63,400—82,400) kg CH_4 hr^{-1}, is not statistically significantly different from a top-down, aircraft-based estimate from the Barnett Coordinated Campaign[21], 76,000±13,000 kg CH_4 hr^{-1}, which quantified emissions in areas intermediate to the 8-county core production area and 25-county region. The bottom-up estimate of thermogenic emissions from O&G and other fossil sources, 48,400 (42,100—56,400) kg CH_4 hr^{-1}, is within the uncertainty bounds of the top-down estimate, 60,000±11,000 kg CH_4 hr^{-1}, of fossil emissions determined from source apportionment based on airborne ethane observations during the campaign.[20,21] Future studies comparing top-down and bottom-up emission estimates should assure that emission inventories rely on comprehensive activity factors and contemporary emission factors that account for the highest emitting sites. Such efforts are likely to results in better agreement between top-down and bottom-up methods than previously has been reported.

Figure 1. Sampled emission rate distributions by O&G sector with superimposed values of Barnett Coordinated campaign measurements. The blue lines are the cumulative distribution functions of sampled distributions used in Monte Carlo simulations, which include production site measurements made using unbiased sampling during the campaign[28] and a national dataset of gathering stations and processing plants[31]. The vertical lines are emission rates of sites measured during the Barnett Campaign using sampling biased towards high emission sites.[27,29,30] The values exceeding the maximum of the sampled distributions are used as fat-tail site distributions in the Monte Carlo simulations.
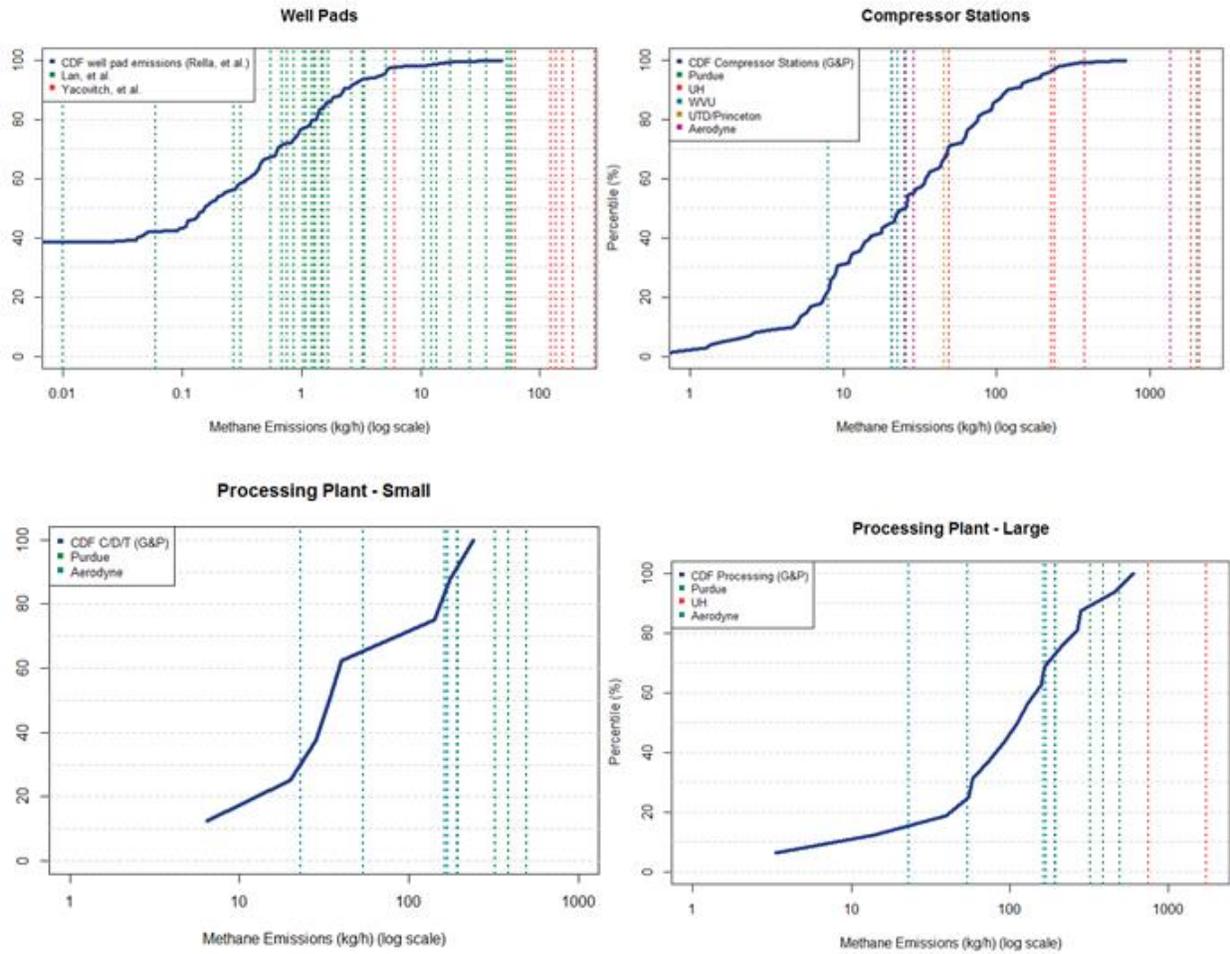
Table 1. Activity factors, methane emissions, and percent of emissions from fat-tail sites by source category for the 25-county Barnett Shale region. Numbers in parentheses are the 95th confidence interval. Estimates assumes a 0.25% fat-tail probability for production sites[15], 1% fat-tail probability for compressor stations, and 2% fat-tail probability for processing plants.

Source	Activity Factor	Emissions (kg CH_4 hr^{-1})	Contribution from Fat-tail sites (%)
Gas Production Sites	15,044 well pads	16,400 (15,400 - 17,300)	11% (8 – 13%)
Oil Production Sites	5,842 well pads	1,800 (1,700 - 1,900)	
Well Completions	38 gas wells 36 oil wells	150 (30 - 290)	
Gathering Compressor Stations	259 facilities	18,700 (12,900 – 26,000)	33% (14 – 51%)
Gathering Pipelines	20,100 miles	940 (760 - 1,200)	
Processing Plants	22 large plants 16 small plants	5,500 (3,700 - 8,100)	11% (4 – 21%)
Transmission & Storage Compressor Stations	17 facilities	1,600 (850 – 1,700)	33% (14 – 51%)
Transmission Pipelines	3,300 miles	230 (190 - 300)	
Local Distribution	5,730,000 people	920 (750 – 1,600)	
O&G Subtotal		**46,200 (40,000 - 54,100)**	19% (14 – 26%)
Abandoned Wells	57,600 wells	630 (320 - 1,300)	
Residential & Commercial End Use	5.6 MMcf/hr gas delivered	160 (30 – 1,600)	
Industrial Facilities	56 facilities	60 (30 - 110)	
Onroad Vehicles (Natural Gas)	0.3 MMcf/hr gas delivered	14 (7 – 68)	
Onroad Vehicles (Gasoline & Diesel)	65 billion vehicle miles traveled/yr	150	
Geological Seepage	57,900 km^2	1,100	
Thermogenic Subtotal		**48,400 (42,100 – 56,400)**	18% (14 – 26%)
Landfills	21 GHGRP landfills 712 other landfills	11,300 (5,000 - 16,900)	
Livestock	980,00 beef cattle 190,00 dairy cattle	11,300 (9,500 - 14,300)	
Wastewater Treatment	5,730,000 people	760 (560 - 670)	
Biogenic Subtotal		**24,000 (17,200 – 30,100)**	
Emissions Total		**72,300 (63,400 - 82,400)**	12% (9 – 15%)
% O&G		64% (52 – 78%)	
%Thermogenic		67% (55 – 81%)	
%Biogenic		33% (23 – 43%)	

Figure 2. Spatially resolved methane emissions of the Barnett Shale region showing total, thermogenic, and biogenic emissions in 4 km x 4 km grid cells. Total emissions are reported in Table 2. The purple line is the boundary of the 25-county Barnett Shale region and the blue line is the boundary of the 8-county core production area.
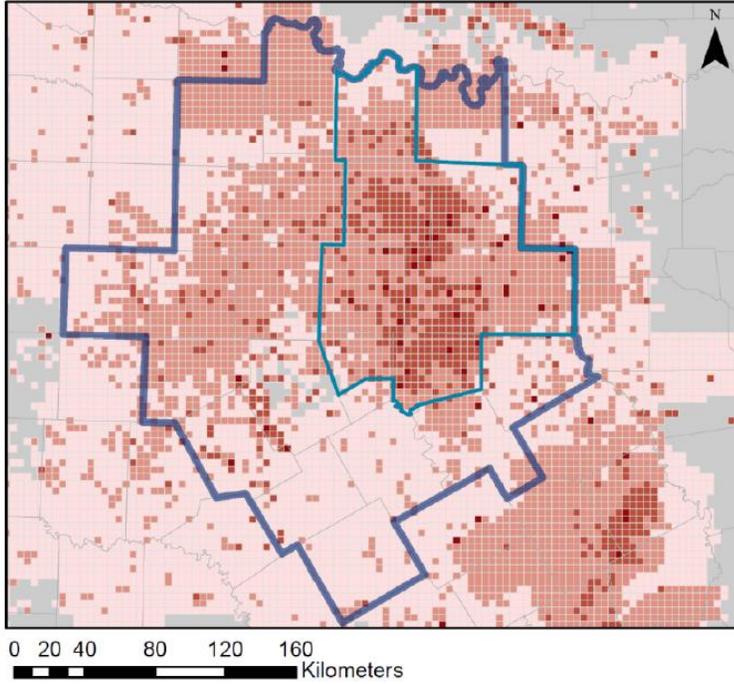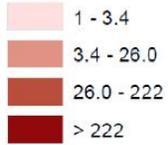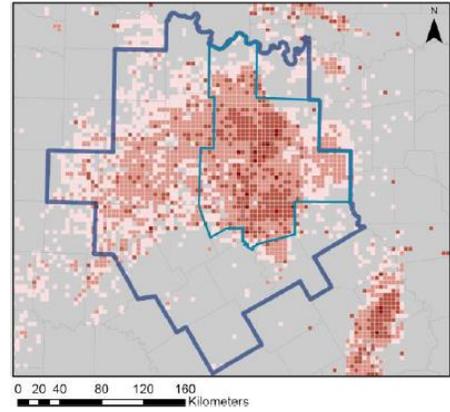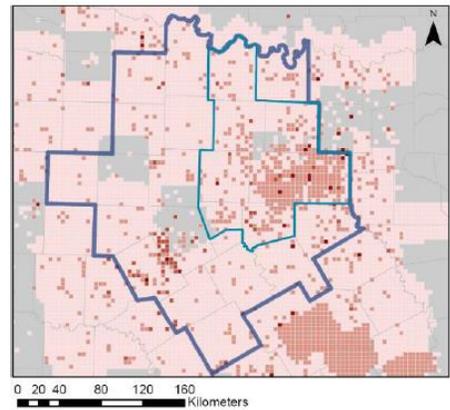
Table 2. A comparison of the 25-county Barnett region O&G methane emission inventories from this paper to alternative inventories constructed from GHGI, GHGRP, and EDGAR.[16-19]

Sector	October 2013 Emissions (kg CH_4 hr^{-1})			
	Lyon et al. (95th percent CI)	GHGI	GHGRP	EDGAR
Production	18,400 (17,100 – 19,500)	12,700	14,550	
Gathering	19,600 (13,700 – 27,200)	2,700	4	
Processing	5,500 (3,700 – 8,100)	8,700	800	
Transmission & Storage	1,800 (1,300 – 2,500)	2,700	300	
Local Distribution	920 (750 – 1,600)	4,300	1,350	
Total O&G	**46,200 (40,000 – 54,100)**	**31,000**	**17,000**	**10,800**

Table 3. A comparison of activity factors (AF) and emissions factors (EF) for the 25-county Barnett region from this paper, GHGI, and GHGRP. GHGI factors are derived from national emissions and activity factors prorated by production, gas processed, and pipeline miles.[17] GHGRP factors are counts and average 2013 emissions of reporting facilities in the Barnett region.[16]

		Lyon et al.	GHGI	GHGRP
Active Oil and Gas wells (excludes completions)	AF	20,900 well pads 39,000 wells	34,800 wells	15,900 wells**
	EF	0.87 (0.82 – 0.92) kg CH_4 hr^{-1} pad^{-1} 0.47 (0.44 – 0.49) kg CH_4 hr^{-1} well^{-1}	0.35 kg CH_4 hr^{-1} well^{-1}	0.91 kg CH_4 hr^{-1} well^{-1}
Gathering Stations (excludes pipelines)	AF	259 stations	1 large station 782 total stations*	54 stations
	EF	72 (50 – 100) kg CH_4 hr^{-1}	1.7 kg CH_4 hr^{-1}*	0.06 kg CH_4 hr^{-1}
Processing Plants	AF	38 plants (16 small & 22 large)	55 plants	22 plants
	EF	average: 145 (84 – 231) kg CH_4 hr^{-1} small: 84 (45 – 133) kg CH_4 hr^{-1} large: 190(112 – 301) kg CH_4 hr^{-1}	159 kg CH_4 hr^{-1}	36 kg CH_4 hr^{-1}
Transmission & Storage Stations (excludes pipelines)	AF	17 stations	25 stations	5 stations
	EF	72 (50 – 100) kg CH_4 hr^{-1}	103 kg CH_4 hr^{-1}	61 kg CH_4 hr^{-1}

*GHGI only includes station counts for large gathering compressor stations. An alternative estimate of total station counts is based on the number of compressor engines. The emission factor is based on this alternative AF. **GHGRP data does not include well counts. Activity data was estimated from the county-level well counts reported in DI Desktop for GHGRP reporting operators.

Acknowledgements

CHAPTER 2 REFERENCES

(1)   Alvarez, R. A; Pacala, S. W.; Winebrake, J. J.; Chameides, W. L.; Hamburg, S. P. Greater focus needed on methane leakage from natural gas infrastructure. *Proc. Natl. Acad. Sci. U. S. A*. **2012**, *109* (17), 6435–6440; DOI 10.1073/pnas.1202407109.

(2)   Allen, D. T.; Torres, V. M.; Thomas, J.; Sullivan, D. W.; Harrison, M.; Hendler, A.; Herndon, S. C.; Kolb, C. E.; Fraser, M. P.; Hill, A. D.; et al. Measurements of methane emissions at natural gas production sites in the United States. *Proc. Natl. Acad. Sci.* **2013**, 110, 17768–17773; DOI 10.1073/pnas.1304880110.

(3)   Brantley, H. L.; Thoma, E. D.; Squier, W. C.; Guven, B. B.; Lyon, D. Assessment of Methane Emissions from Oil and Gas Production Pads using Mobile Measurements. *Environ. Sci. Technol.* **2015**, 49, 3219-3227; DOI: 10.1021/es5052809.

(4)   Pétron, G.; Karion, A.; Sweeney, C.; Miller, B. R.; Montzka, S. A.; Frost, G. J.; Trainer, M.; Tans, P.; Andrews, A.; Kofler, J.; et al. A new look at methane and nonmethane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin. *J. Geophys. Res. Atmos.* **2014**, 119, 6836–6852; DOI 10.1002/2013JD021272.

(5)   Wecht, K. J.; Jacob, D. J.; Frankenberg, C.; Jiang, Z.; Blake, D. R. Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data. J*. Geophys. Res. Atmos.* **2014**, 119, 7741–7756; DOI 10.1002/2014JD021551.

(6)   Caulton, D. R.; Shepson, P. B.; Santoro, R. L.; Sparks, J. P.; Howarth, R. W.; Ingraffea, A. R.; Cambaliza, M. O. L.; Sweeney, C.; Karion, A.; Davis, K. J.; et al. Toward a better understanding and quantification of methane emissions from shale gas development. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, 111, 6237–6242; DOI 10.1073/pnas.1316546111.

(7)   Miller, S. M.; Wofsy, S. C.; Michalak, A. M.; Kort, E. a; Andrews, A. E.; Biraud, S. C.; Dlugokencky, E. J.; Eluszkiewicz, J.; Fischer, M. L.; Janssens-Maenhout, G.; et al. Anthropogenic emissions of methane in the United States. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, 110, 20018–20022; DOI 10.1073/pnas.1314392110.

(8)   Brandt, A. R.; Heath, G. A.; Kort, E. A.; O'Sullivan, F.; Pétron, G.; Jordaan, S. M.; Tans, P.; Wilcox, J.; Gopstein, A. M.; Arent, D.; et al. Methane Leaks from North American Natural Gas Systems. *Science.* **2014,** 343 (6172), 733–735; DOI 10.1126/science.1247045.

(9)   Allen, D. T. Methane emissions from natural gas production and use: reconciling bottom-up and top-down measurements. *Curr. Opin. Chem. Eng.* **2014**, 5, 78–83; DOI 10.1016/j.coche.2014.05.004.

(10)   Texas Commission on Environmental Quality. Barnett Shale Information; TCEQ: Austin, TX, 2014; http://www.rrc.state.tx.us/oil-gas/major-oil-gas-formations/barnett-shale-information/

(11)   Browning, J.; Ikonnikova, S.; Gülen, G.; Tinker, S. Barnett Shale Production Outlook. *SPE Econ. Manag.* **2013**, 5, 89–104.

(12)   Zavala-Araiza, D.; Sullivan, D. W.; Allen, D. T. Atmospheric hydrocarbon emissions and concentrations in the Barnett shale natural gas production region. *Environ. Sci. Technol.* **2014**, 48, 5314–5321; DOI 10.1021/es405770h.

(13)   Texas Commission on Environmental Quality. Barnett Shale Area Special Inventory, Phase 2; TCEQ: Austin, TX, 2014; https://www.tceq.texas.gov/assets/public/implementation/ air/ie/pseiforms/ Barnett%20Shale%20Area%20Special%20Inventory.pdf

(14)   Eastern Research Group. *City of Fort Worth Natural Gas Air Quality Study*; City of Fort Worth: Fort Worth, TX, 2011; http://fortworthtexas.gov/gaswells/air-quality-study/final/

(15)   Zavala-Araiza, D.; Lyon, D.; Alvarez, R. A.; Palacios, V.; Lan, X.; Talbot, R.; Allen, D. T.; Hamburg, S. P. Towards a Functional Definition of Methane Super-Emitters: Application to Natural Gas Production Sites. *Environ. Sci. Technol., Submitted.*

(16)   United States Environmental Protection Agency. Greenhouse Gas Reporting Program; EPA: Washington, DC, 2014; http://ghgdata.epa.gov/ghgp/main.do (accessed Nov 1, 2014).

(17)   United States Environmental Protection Agency. Draft Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2013; EPA: Washington, DC, 2015; http://www.epa.gov/ climatechange/ ghgemissions/ usinventoryreport.html

(18)   European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL). Emission Database for Global Atmospheric Research, release version 4.2; 2011; http://edgar.jrc.ec.europa.eu

(19)   Kort, E. A.; Frankenberg, C.; Costigan, K. R.; Lindenmaier, R.; Dubey, M. K.; Wunch, D. Four corners: The largest US methane anomaly viewed from space. *Geophys. Res. Lett.* **2014**, 41, 6898–6903. DOI 10.1002/2014GL061503.

(20)   Karion, A.; Sweeney, C.; Kort, E. A.; Shepson, P. B.; Brewer, A.; Cambaliza, M. O. L.; Conley, S.; Davis, K. J.; Deng, A.; Hardesty, M.; et al. Aircraft-based estimate of total methane emissions from the Barnett Shale region. *Environ. Sci. Technol., Submitted.*

(21)   Smith, M. L.; Kort, E. A.; Karion, A.; Sweeney, C.; Herndon, S. C.; Yacovitch, T. I. Airborne ethane observations in the Barnett Shale: Quantification of ethane flux and attribution of methane emissions. *Environ. Sci. Technol., Submitted.*

(22)   Roy, A. A.; Adams, P. J.; Robinson, A. L. Air pollutant emissions from the development, production, and processing of Marcellus Shale natural gas. *J. Air Waste Manage. Assoc.* **2013**, 64, 19–37; DOI 10.1080/10962247.2013.826151.

(23)   Jeong, S.; Millstein, D.; Fischer, M. L. Spatially explicit methane emissions from petroleum production and the natural gas system in California. *Environ. Sci. Technol.* **2014**, 48, 5982–5990; DOI 10.1021/es4046692.

(24)   Texas Commission on Environmental Quality. Comprehensive Air Quality Model Domain Shapefile TX 4 km; TCEQ: Austin, TX, 2014; ftp://amdaftp.tceq.texas.gov/pub/TX/map/

(25)   Texas Commission on Environmental Quality. Air Permit Database; TCEQ: Austin, TX, 2014; http://www2.tceq.texas.gov/airperm/

(26)   Drillinginfo. DI Desktop; Drillinginfo: Austin, TX, 2015. http://www.didesktop.com/ (accessed March 1, 2015).

(27)   Yacovitch, T. I.; Herndon, S. C.; Pétron, G.; Kofler, J.; Lyon, D.; Zahniser, M. S.; Kolb, C. E. Mobile Laboratory Observations of Methane Emissions in the Barnett. *Environ. Sci. Technol.* **2015**; DOI 10.1021/es506352j.

(28)   Rella, C. W.; Tsai, T.; Botkin, C.; Crosson, E.; Steele, D. Measuring Emissions from Oil and Natural Gas Producing Well Pads in the Barnett Shale Region Using the Novel Mobile Flux Plane Technique. *Environ. Sci. Technol.* 2015, 49, 4742-4748; DOI 10.1021/acs.est.5b00099.

(29)   Lan, X.; Talbot, R.; Laine, P.; Torres, A. Characterizing Fugitive Methane Emissions in the Barnett Shale Area Using a Mobile Laboratory. *Environ. Sci. Technol., Submitted*.

(30)   Lavoie, T. N.; Shepson, P. B.; Cambaliza, M. O. L.; Karion, A.; Sweeney, C.; Hirst, B.; Yacovitch, T. I.; Herndon, S. C.; Lan, X.; Lyon, D.; et al. Aircraft-Based Measurements of Point Source Methane Emissions in the Barnett Shale Basin. *Environ. Sci. Technol., Submitted*.

(31)   Mitchell, A. L.; Tkacik, D. S.; Roscioli, J. R.; Herndon, S. C.; Yacovitch, T. I.; Martinez, D. M.; Vaughn, T. L.; Williams, L.; Sullivan, M.; Floerchinger, C.; et al. Measurements of methane emissions from natural gas gathering facilities and processing plants: Part 2. measurement results. *Environ. Sci. Technol.* **2015**, 49, 3219-3227; DOI 10.1021/es5052809.

(32)   Standards of Performance for New Stationary Sources: Subpart OOOO—Standards of Performance for Crude Oil and Natural Gas Production, Transmission and Distribution. Public Law 77 FR 49542, 2012. http://www.ecfr.gov/cgi-bin/text-idx?SID=true&node=sp40.7.60.oooo

(33)   Lamb, B. K.; Edburg, S. L.; Ferrara, T. W.; Howard, T; Harrison, M. R.; Kolb, C. E.; Townsend-Small, A.; Dyck, W; Possolo, A; Whetstone, J. R. Direct Measurements Show Decreasing Methane Emissions from Natural Gas Local Distribution Systems in the United States. *Environ. Sci. Technol.,* **2015;** DOI 10.1021/es505116p.

(34)   Kang, M.; Kanno, C. M.; Reid, M. C.; Zhang, X.; Mauzerall, D. L.; Celia, M. A.; Chen, Y.; Onstott, T. C. Direct measurements of methane emissions from abandoned oil and gas wells in Pennsylvania. *Proc. Natl. Acad. Sci.* **2014**, 111, 201408315; DOI 10.1073/pnas.1408315111.

(35)   Eastern Research Group. Emission Factor Documentation for AP-42 Section 1.4 Natural Gas Consumption; EPA: Washington, DC, 1998; http://www.epa.gov/ttnchie1/ap42/ ch01/ bgdocs/b01s04.pdf

(36)   Energy Information Administration. Natural Gas Consumption by End Use; EIA: Washington, DC, 2014; http://www.eia.gov/dnav/ng/ng_cons_sum_a_EPG0_vgt_mmcf_m.htm

(37)   United States Census Bureau. 2010 Census Summary File 1; US Census: Washington, DC, 2011; http://www2.census.gov/census_2010/04-Summary_File_1/Texas/

(38)   Walpert, T. H. Measuring Residential Natural Gas Leakage and its Impact on CH4 Emission Inventories in California. Senior Thesis, University of California-Berkeley, Berkeley, CA; 2013; http://nature.berkeley.edu/classes/es196/projects/2013final/WalpertT_2013.pdf

(39)   McKain, K.; Down, A.; Raciti, S. M.; Budney, J.; Hutyra, L. R.; Floerchinger, C.; Herndon, S.; Nehrkorn, T.; Zahniser, M. S.; Jackson, R. B.; et al. Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. *Proc. Nat. Acad. Sci. U. S. A.* **2015**, 112, 1941-1946; DOI 10.1073/pnas.1416261112.

(40)   United States Environmental Protection Agency. 2011 National Emissions Inventory; EPA: Washington, DC, 2013; http://www.epa.gov/ttn/chief/net/2011inventory.html

(41)   United States Environmental Protection Agency. Motor Vehicle Emission Simulator (MOVES) Version 2010b; EPA: Washington, DC, 2010; http://www.epa.gov/otaq/models/moves/moves-docum.htm

(42)   Texas Commission on Environmental Quality. Air Modeling FTP site, MOVES output files; TCEQ: Austin, TX, 2014; ftp://amdaftp.tceq.texas.gov/pub/Mobile_EI/Trends/mvs/

(43)   Etiope, G.; Klusman, R. W. Geologic emissions of methane to the atmosphere. *Chemosphere*. **2002**, 49, 777–789; DOI 10.1016/S0045-6535(02)00380-6.

(44)   Texas Commission on Environmental Quality. Inventory of Closed Municipal Solid Waste Landfills; TCEQ: Austin, TX, 2014; http://www.tceq.state.tx.us/permitting/waste_permits/ waste_planning/ wp_closed_lf_inv.html

(45)   United States Environmental Protection Agency. Technical Support Document for the Landfill Sector: Proposed Rule for Mandatory Reporting of Greenhouse Gases; EPA: Washington, DC, 2009; http://www.epa.gov/ghgreporting/documents/pdf/archived/tsd/TSD-Landfill-sector.pdf

(46)   Texas Commission on Environmental Quality. Water Quality General Permits & Registration Search; TCEQ: Austin, TX, 2014; http://www2.tceq.texas.gov/wq_dpa/index.cfm

(47)   National Agriculture Statistics Service. Texas Cattle Inventory by County; United States Department of Agriculture: Washington, DC, 2014; http://www.nass.usda.gov/ Statistics_by_State/Texas/Publications/County_Estimates/ce_tables/cecatt1.htm

(48)   Dong, H.; Mangino, J.; McAllister, T. A.; Hatfield, J. L.; Johnson, D. E.; Lassey, K. R.; de Lima, M. A.; Romanovskaya, A. 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Volume 4: Agriculture, Forestry, and Other Land Use, Chapter 10: Emissions from Livestock and Manure Management; IPCC: Paris, France, 2006; http://www.ipcc-nggip.iges.or.jp/public/ 2006gl/ pdf/4_Volume4/V4_10_Ch10_Livestock.pdf

(49)   United States Census Bureau. 2013 TIGER/Line® Shapefiles: Census Tracts; US Census: Washington, DC, 2013; https://www.census.gov/cgi-bin/geo/shapefiles2013/layers.cgi

(50)   Allen, D. T.; Sullivan, D. W.; Zavala-Araiza, D.; Pacsi, A. P.; Harrison, M.; Keen, K.; Fraser, M. P.; Daniel Hill, A.; Lamb, B. K.; Sawyer, R. F.; et al. Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Liquid Unloadings. *Environ. Sci. Technol.* **2014**, 49, 641-648; DOI 10.1021/es504016r.

(51)   Subramanian, R.; Williams, L.; Vaughn, T. L.; Zimmerle, D.; Roscioli, J. R.; Herndon, S. C.; Yacovitch, T. I.; Floerchinger, C.; Tkacik, D. S.; Mitchell, A. L.; et al. Methane emissions from natural gas compressor stations in the transmission and storage sector: Measurements and comparisons with the EPA greenhouse gas reporting program. *Environ. Sci. Technol.* **2015**, DOI 10.1021/es5060258.

(52)   Zavala-Araiza, D.; Allen, D. T.; Harrison, M.; George, F. C.; Jersey, G. B. Allocating Methane Emissions to Natural Gas and Oil Production in Shale Formations. *Environ. Sci. Technol.* **2015**,3, 492-498; DOI 10.1021/sc500730x.

(53)   Harrison, M. R.; Shires, T.; Wessels, J. K.; Cowgill, R. M. Methane emissions from the natural gas industry; EPA: Washington, DC,1996; http://www.epa.gov/methane/gasstar/tools/related.html

(54)   Mandatory Greenhouse Gas Reporting: Subpart W - Petroleum and Natural Gas Systems. Public Law 75 FR 79140, 2010; http://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID= d7a88dd57a772702658ef74704f9f395&n= pt40.21.98&r=PART&ty=HTML#sp40.21.98.w

(55)   Mandatory Greenhouse Gas Reporting: Subpart C - General Stationary Fuel Combustion Sources. Public Law 78 FR 71952, 2013; http://www.ecfr.gov/cgi-bin/text-idx?SID=d7a88dd57a772702658ef74704f9f395&node= sp40.21.98.c

(56)   United States Environmental Protection Agency. AP-42, Fifth Edition, Volume I, Chapter 3: Stationary Internal Combustion Sources, Part 3.2: Natural Gas-fired Reciprocating Engines; EPA: Washington, DC, 2000; http://www.epa.gov/ttn/chief/ap42/ch03/final/c03s02.pdf

(57)   Johnson, D.; Covington, A.; Clark, N. Methane emissions from leak and loss audits of natural gas compressor stations and storage facilities. *Environ. Sci. Technol. Submitted*.

(58)   40 CFR Part 98 Greenhouse Gas Reporting Rule: 2015 Revisions and Confidentiality Determinations for Petroleum and Natural Gas Systems; Proposed Public Law, 2014; http://www.gpo.gov/fdsys/pkg/FR-2014-12-09/pdf/2014-28395.pdf

CHAPTER 2 APPENDIX

S1. Activity factor estimation methods

The number and location of O&G and other methane-emitting facilities were compiled from multiple state and federal databases. Facilities with annual greenhouse gas emissions of ≥25,000 metric tons carbon dioxide equivalents (CO_2e) are required to submit data including site location to the United Stations Environmental Protection Agency (EPA) Greenhouse Gas Reporting Program (GHGRP)[1]. Point source facilities reporting methane emissions in 2013 were classified into source categories based on which subpart they report under: natural gas transmission (Subpart W-NGTC)[2], natural gas processing (Subpart W-PROC)[2], landfills (Subpart HH)[3], and other industrial (all other subparts). Natural gas gathering facilities currently are exempt from GHGRP Subpart W but report stationary combustion emissions (Subpart C)[4] if they meet the reporting threshold. Facilities reporting for stationary combustion only were classified as natural gas gathering if the facility name and North American Industrial Classification System code suggested it was an O&G facility (211111-Crude petroleum and natural gas extraction, 211112-Natural gas liquid extraction, and 486210-Pipeline transportation of natural gas); otherwise, they were classified as non-O&G industrial sources.

Additional O&G sources were identified based on data from the Texas Commission on Environmental Quality (TCEQ) 2009 Barnett Shale Area Special Inventory (BSASI)[5] and air permit database[6]. TCEQ required all companies with upstream and midstream facilities either producing or handling gas from the Barnett Shale in 2009 to submit data including site location and an equipment census .[5] The BSASI data was filtered to exclude well pads by removing all upstream facilities and the subset of midstream facilities without engines. The midstream facilities with engines were classified as a compressor station, storage facility, or processing plant based on the facility name. In general, facilities with names including "station", "compression", or "CS" were classified as compressor stations; "processing" or "plant" as processing plants; and "storage" as storage facilities. A list of facilities with TCEQ engine permits compiled in 2014 by ICF International for Clean Air Task Force was used to identify facilities not included in the BSASI. Additionally, the TCEQ permit database was queried to identify facilities with O&G general permits

missing from the other databases.[6] The additional permitted facilities were classified into site types based on name using a similar approach as the BSASI facilities. The coordinates of these facilities were determined by using Google Earth to match the driving directions provided by the TCEQ permit data to a facility visible in satellite imagery. For most of the region, Google Earth imagery was from April 2013 or October 2013.

Compressor stations not reporting to the GHGRP were classified as gathering or transmission based on their proximity to gathering and transmission pipelines. GIS shapefiles of gathering and transmission pipelines were obtained from DI Desktop.[7] ArcGIS (Esri) was used to create a 1.5 km buffer around the natural gas transmission pipelines. All the compressor stations outside the buffer were classified as gathering. For compressors stations that were inside the buffer, a manual classification scheme was performed, using Google Earth imagery in conjunction with natural gas gathering and transmission pipeline maps. This method minimizes misclassification as transmission of facilities near transmission pipelines with no connection to the facility, but there remains some uncertainty in the classification.

The multiple databases included several instances of facilities with similar names and different locations or different names and similar locations. Facility locations were quality controlled to remove duplicate facilities and correct inaccurate coordinates by checking Google Earth imagery. Contiguous facilities were considered individual sites if they were operated by different companies and/or divided by a fence line. Facilities were also excluded if Google Earth imagery indicated they were decommissioned (e.g., all engines had been removed).

Emissions from O&G production were calculated at the production site level, since that is the spatial scale at which measurements were made. DI Desktop, a commercial database of well data compiled from state regulatory agencies, was used to obtain the coordinates of O&G wells in the Barnett region with production in October 2013.[7] Individual wells were grouped into production sites by aggregating wells within a 100 m radius into a single site as described in Zavala-Araiza et al.[8]

S2. Production site Monte Carlo simulation methods

Production site emissions were estimated using a more complex approach that defined fat-tail sites based on based on proportional loss rates (methane emitted relative to methane produced). The method is briefly described below and in detail in Zavala-Araiza et al.[8] Activity factors were based on estimated O&G production site counts. Emission factors were derived with Monte Carlo simulations drawing from site emission rate distributions constructed using data from the Barnett Coordinated Campaign. Rella et al.[9] reports emissions from 186 production sites that were selected from sites less than 150 meters upwind of public roads in nine Barnett counties. Additional emission estimates of 40 production sites selected by targeted sampling were reported in Lan et al.[10] and Yacovitch et al.[11] Activity and measurement data were classified into four cohorts based on gas production. Production site emissions as a percentage of gas production were calculated using production data from DI Desktop[7]. Production site measurements were further classified into three categories based on their percentile of production-normalized emissions specific to each production class: alpha (α) sites (0—85 percentile), beta (β) sites (≥85 percentile), and gamma (γ) sites (≥ maximum value of Rella et al. dataset). As described in Zavala-Araiza et al., a fat-tail probability of γ-sites of 0.25% was chosen and a sensitivity analysis was performed to test the effect of differing probabilities on estimated emissions. Zavala-Araiza et al. reports emissions only for gas-producing sites. For this paper, gas-producing site emissions were divided into gas sites and oil sites based on the well type reported in DI Desktop. In addition, emissions from oil sites with no gas production were estimated using a methane emission factor of 5.14×10^{-3} kg CH_4 hr^{-1} $well^{-1}$ based on the Petroleum Systems stripper well emission factor in the GHGI.[12]

S3. Well completion emission estimation methods

Completion flowback emissions, which occur when a well is vented after hydraulic fracturing to prepare for routine production, were estimated for individual well completions based on data obtained from DI Desktop.[7] For 73 wells in the region with reported completion dates from October 14—30, 2013, data were compiled including well type (gas or oil), latitude/longitude, and an estimate of initial gas production based on average daily production during the second month of production. Potential completion flowback emissions were assumed to equal initial production adjusted by an assumed gas composition of 78.8%

methane. Well completions were assumed to emit at a constant rate for 3 days, which is a conservatively low estimate based on the 85-hour average completion flowback duration reported in Allen et al.[13] In comparison, O'Sullivan and Paltsev assumed potential completion emissions increase linearly from zero to initial production over a 9 day period[14]; our approach results in about twice the average emission rate but half the total emissions of these alternative assumptions. At the time of the campaign, gas wells were required by the EPA New Source Performance Standard Subpart OOOO to control completion emissions by flaring or sending gas to sales.[15] Therefore, gas well completions were assumed to be controlled with 98% control efficiency. This high control efficiency is supported by Allen et al., who determined actual emissions of completion flowbacks with control or capture were 1% of potential emissions. [13] Oil wells, which are not required to control or capture gas, were assumed to be uncontrolled and emit 100% of potential emissions. Emissions were aggregated on a daily basis for all ongoing completion events. The average of the daily emission estimates during the Barnett Coordinated campaign was used as the central estimate for the emission inventory. The minimum and maximum daily estimates were used as the lower and upper bound estimates.

S4. Local distribution emission estimation methods

Natural gas distribution emissions were estimated using data from a recent national study of methane emissions from local distribution pipelines and metering and regulating (M&R) stations (Lamb et al.).[16] Activity factors were based on 2013 data reported by Atmos Energy Corporation, the utility serving the vast majority of customers in the Barnett region. Atmos Mid-Texas distribution main pipeline miles, number of services by type, and number of leaks repaired were obtained from the Pipeline and Hazardous Materials Safety Administration.[17] Emissions were estimated using the national average equivalent leak ratio (1.6 leaks per leaks repaired) and type-specific main and service emission factors from Lamb et al.[16] Atmos Texas activity factors for above ground metering stations and transfer stations were obtained from the GHGRP.[1] M&R activity factors were disaggregated by type by assuming that above ground metering stations were distributed in the same proportion as national activity factors from the EPA Greenhouse Gas Inventory (GHGI), the ratio of below ground to above ground stations was the same as the GHGI, and all transfer stations were M&R stations with > 300 pound per square inch (psi)

pressure.[12] Emissions were estimated using type-specific station emission factors from Lamb et al.[16]

Since the Atmos Mid-Texas and Atmos Texas regions are larger than the 25-county Barnett Shale region, emissions were multiplied by the ratio of customer meters in the Barnett region to the Atmos service areas. The number of customer meters in Atmos service areas and major cities were obtained from the Texas Railroad Commission[18,20]; meters in the Barnett region were estimated by scaling up the number of Atmos residential customer meters in the cities of Dallas and Fort Worth[18] by population[19] and the ratio of commercial and industrial to residential meters in the Atmos Mid-Texas service area[20]. Emissions from customer meters and maintenance/upsets, which are not measured in Lamb et al., were estimated using GHGI national emissions prorated by customer meters and pipeline miles, respectively[12]. The upper confidence limit uncertainty (+71%) was based on the emission factor uncertainties of Lamb et al.[16]; for the lower confidence limit, EPA's uncertainty estimate (-19%) for GHGI Natural Gas Systems was used since Lamb et al. only reports upper confidence limits.[12]

S5. Alternative emission inventory methods

The GHGI based emission inventory was constructed by prorating 2013 national net emissions by source category from the GHGI Natural Gas Systems and Petroleum Systems sectors.[12] Natural gas production emissions were prorated by the ratio of Barnett region and national 2013 gas production.[7,21] Natural gas gathering is grouped within the natural gas production sector in the GHGI.[12] We assigned pipeline leaks, pipeline blowdowns, gas engines, small reciprocating compressors, large reciprocating compressors, compressor starts, and compressor blowdowns to the gathering sector. Natural gas gathering emissions were prorated by 2013 gas production.[7,21] Natural gas processing emissions were prorated by 2013 gas processed, which was estimated by adjusting the Texas fraction of national gas processed[22] by the fraction of Texas gas production in the Barnett region[7]. Natural gas transmission and storage emissions were prorated by 2013 transmission pipeline miles.[23] Natural gas distribution emissions were prorated by 2013 gas delivered to customers excluding electric power, which was estimated by adjusting the Texas fraction of national gas delivered[19] by the fraction of Texas population in the Barnett region[19]. Petroleum Systems production emissions were prorated by the ratio of Barnett region and national 2013 oil production.[7,24]

The GHGRP based inventory was constructed by aggregating GHGRP reported 2013 emissions from facilities in the 25-county Barnett region.[1] The classification of GHGRP facilities is described in section S1. For all facilities except onshore oil and gas production and natural gas distribution, quality assured facility latitude/longitude was used to determine which sites were in the Barnett region; several of the reported locations were inaccurate and corrected as described in section S1. GHGRP local distribution facilities are defined as a company's statewide operations. Reported emissions from Atmos Energy Corporation – Texas, the utility serving the vast majority of the Barnett region, were adjusted by the estimated fraction (0.542) of their statewide customer meters in the Barnett region (described in section S4.) GHGRP onshore production facilities are defined as all of a company's production sites in a geologic basin. The Barnett region is comprised of five basins including three that extend outside the Barnett. County-level 2013 gas and oil production were used to determine the fraction of energy production in each basin contained in the 25-county Barnett region: 350 – South Oklahoma Belt (0.143), 400 – Ouachita Folded Belt (0.549), 415 – Strawn Basin (1.0), 420 – Fort Worth Syncline (1.0), and 425 – Bend Arch (0.854).[7] Emissions from well completions and workovers, atmospheric tanks, and well venting for liquids unloading are reported at the county-level; for these sources, emissions were included only from the Barnett region counties. Emissions from all other sources, which are reported at the basin-level, were adjusted by the fraction of basin energy production in the Barnett region.

Figure SI1. A map of spatially referenced activity factors of major methane sources in the Barnett Shale region. Facilities are color-coded by type (red = compressor station, blue = processing, yellow = landfill, green = confined animal feeding operation). The orange line is the boundary of the 25-county region. The blue line is the boundary of the 8-county core production area.
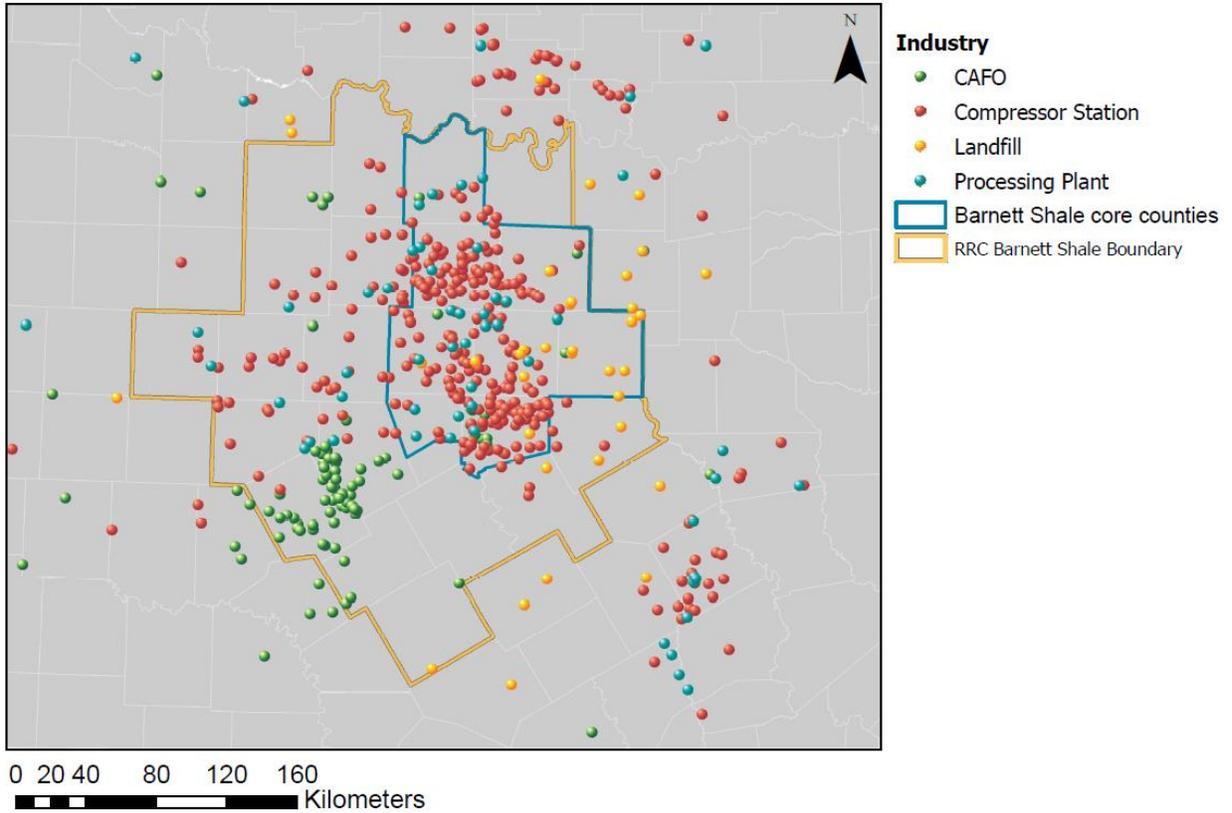
Figure SI2. Illustration of Monte Carlo simulation methods

Sampled Distribution

kg CH4 hr-1

Fat-tail Site Distribution

kg CH4 hr-1

probability = 1 – fat-tail frequency

[example = 0.98]

probability = fat-tail frequency

[example = 0.02]

random selection with replacement from one of two
distributions for every site in 25 county region
[example = 100 sites]

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

Σ

10,000 iterations

Regional Emissions Distribution

97.5%

Mean

2.5%

kg CH4 hr-1

Divide mean, 2.5th percentile, and 97.5th percentile regional emissions
by number of sites in 25 county region (example = 100 sites)

95th confidence interval emission factor

[example = 343 (228 – 470) kg CH$_4$ hr^{-1} site^{-1}]

[mean emission factor 25% higher from fat-tail sites]

Figure SI3. Sensitivity of O&G facility emission factors to the Monte Carlo simulation probabilities of selecting from the fat-tail site distribution of Barnett Campaign measurements that exceed the maximum emission rate of the sampled distribution based on a national dataset of gathering stations and processing plants. The red line is the central estimate and the blue area is the 95th percent confidence interval.
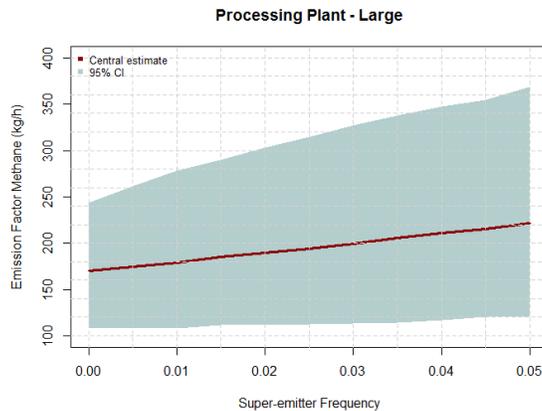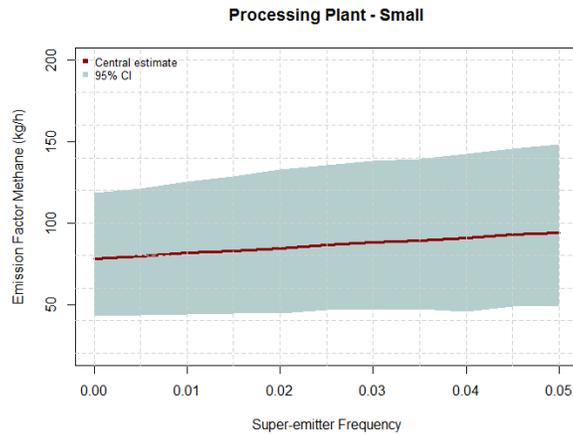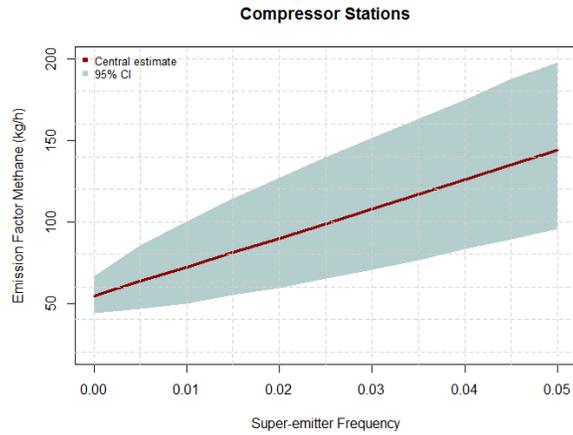
S7. Excel file (es506359c_si_002.xlsx) with O&G facility locations, Monte Carlo emission rate distributions, local distribution emission calculations, and GHGI and GHGRP alternative inventory calculations

Lyon et al_SI_S7.xlsx includes 12 worksheets described below.

Worksheet	Description
S7-1	location of well pads used in Production site Monte Carlo simulations of Zavala-Araiza et al.
S7-2	location of compressor stations and processing plants used in O&G facility Monte Carlo simulations
S7-3	sampled and fat-tail distributions used in small processing plant, large processing plant, and compressor station Monte Carlo simulations
S7-4	comparison of GHG Inventory and GHG Reporting Program based O&G emissions by sector
S7-5	EPA 2013 GHG Inventory emissions by source
S7-6	data used to prorate GHG Inventory national emissions to Barnett Shale region
S7-7	GHGRP O&G production facility emissions for sources reported at county-level
S7-8	GHGRP O&G production facility emissions for sources reported at basin-level
S7-9	GHGRP O&G production emissions adjusted to Barnett Shale region
S7-10	GHGRP emissions from facilities reporting at site-level
S7-11	calculation of GHGI-based activity factors and emission factors for the Barnett Shale region
S7-12	calculation of local distribution emissions

CHAPTER 2 APPENDIX REFERENCES

(1)   United States Environmental Protection Agency. Greenhouse Gas Reporting Program; EPA: Washington, DC, 2014; http://ghgdata.epa.gov/ghgp/main.do (accessed Nov 1, 2014).

(2)   Mandatory Greenhouse Gas Reporting: Subpart W - Petroleum and Natural Gas Systems. Public Law 75 FR 79140, 2010; http://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID= d7a88dd57a772702658ef74704f9f395&n= pt40.21.98&r=PART&ty=HTML#sp40.21.98.w

(3)   Mandatory Greenhouse Gas Reporting: Subpart HH - Municipal Solid Waste Landfills. Public Law 78 FR 71968, 2013; http://www.ecfr.gov/cgi-bin/text-idx?SID=d7a88dd57a772702658ef74704f9f395&node =sp40.21.98.hh&rgn=div6

(4)   Mandatory Greenhouse Gas Reporting: Subpart C - General Stationary Fuel Combustion Sources. Public Law 78 FR 71952, 2013; http://www.ecfr.gov/cgi-bin/text-idx?SID=d7a88dd57a772702658ef74704f9f395&node= sp40.21.98.c

(5)   Texas Commission on Environmental Quality. Barnett Shale Area Special Inventory, Phase 2; TCEQ: Austin, TX, 2014; https://www.tceq.texas.gov/assets/public/implementation/ air/ie/pseiforms/ Barnett%20Shale%20Area%20Special%20Inventory.pdf

(6)   Texas Commission on Environmental Quality. Air Permit Database; TCEQ: Austin, TX, 2014; http://www2.tceq.texas.gov/airperm/.

(7)   Drillinginfo. DI Desktop; Drillinginfo: Austin, TX, 2015. http://www.didesktop.com/ (accessed March 1, 2015.

(8)   Zavala-Araiza, D.; Lyon, D.; Alvarez, R. A.; Palacios, V.; Harriss, R.; Lan, X.; Talbot, R.; Hamburg, S. P. Towards a Functional Definition of Methane Super-Emitters: Application to Natural Gas Production Sites. *Environ. Sci. Technol.,* **2015;** DOI: 10.1021/acs.est.5b00133.

(9)   Rella, C. W.; Tsai, T.; Botkin, C.; Crosson, E.; Steele, D. Measuring Emissions from Oil and Natural Gas Well Pads Using the Novel Mobile Flux Plane Technique. *Environ. Sci. Technol.* **2015**, 49, 4742-4748; DOI 10.1021/acs.est.5b00099.

(10)   Lan, X.; Talbot, R.; Laine, P.; Torres, A. Characterizing Fugitive Methane Emissions in the Barnett Shale Area Using a Mobile Laboratory. *Environ. Sci. Technol.,* **2015**; DOI: 10.1021/es5063055.

(11)   Yacovitch, T. I.; Herndon, S. C.; Pétron, G.; Kofler, J.; Lyon, D.; Zahniser, M. S.; Kolb, C. E. Mobile Laboratory Observations of Methane Emissions in the Barnett. *Environ. Sci. Technol.* **2015**; DOI 10.1021/es506352j.

(12)   United States Environmental Protection Agency. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2013; EPA: Washington, DC, 2015; http://www.epa.gov/ climatechange/ ghgemissions/ usinventoryreport.html

(13)   Allen, D. T.; Torres, V. M.; Thomas, J.; Sullivan, D. W.; Harrison, M.; Hendler, A.; Herndon, S. C.; Kolb, C. E.; Fraser, M. P.; Hill, A. D.; et al. Measurements of methane emissions at natural gas production sites in the United States. *Proc. Natl. Acad. Sci.* **2013**, 110, 17768–17773; DOI 10.1073/pnas.1304880110.

(14) O'Sullivan, F.; Paltsev, S. Shale gas production: potential versus actual greenhouse gas emissions. *Environ. Res. Lett.* **2012**, 7, 044030; DOI 10.1088/1748-9326/7/4/044030.

(15) Standards of Performance for New Stationary Sources: Subpart OOOO—Standards of Performance for Crude Oil and Natural Gas Production, Transmission and Distribution. Public Law 77 FR 49542, 2012. http://www.ecfr.gov/cgi-bin/text-idx?SID=true&node=sp40.7.60.oooo

(16) Lamb, B. K.; Edburg, S. L.; Ferrara, T. W.; Howard, T; Harrison, M. R.; Kolb, C. E.; Townsend-Small, A.; Dyck, W; Possolo, A; Whetstone, J. R. Direct Measurements Show Decreasing Methane Emissions from Natural Gas Local Distribution Systems in the United States. *Environ. Sci. Technol.,* **2015;** DOI 10.1021/es505116p.

(17) Pipeline and Hazardous Materials Safety Administration. Gas Distribution Annual Data – 2010 to present; PHMSA: Washington, DC, 2015; http://phmsa.dot.gov/staticfiles/PHMSA/ DownloadableFiles/Pipeline2data/annual_gas_distribution_2010_present.zip

(18) Texas Railroad Commission. Twenty-five Texas Cities Annual Average Natural Gas Consumption Residential; RRC: Austin, TX, 2013; http://www.rrc.state.tx.us/media/27231/25-avg-ann-con-res-2014.pdf

(19) United States Census Bureau. 2010 Census Summary File 1; US Census: Washington, DC, 2011; http://www2.census.gov/census_2010/04-Summary_File_1/Texas/

(20) Texas Railroad Commission. Investor-Owned Distribution Systems Number of Consumers; RRC: Austin, TX, 2013; http://www.rrc.state.tx.us/media/27246/table-4a.pdf

(21) Energy Information Administration. Natural Gas Gross Withdrawals and Production; EIA: Washington, DC, 2014; http://www.eia.gov/dnav/ng/ng_prod_sum_a_EPG0_FGW_mmcf_a.htm

(22) Energy Information Administration. Natural Gas Plant Processing; EIA: Washington, DC, 2014; http://www.eia.gov/dnav/ng/ng_prod_pp_a_EPG0_ygp_mmcf_a.htm

(23) Pipeline and Hazardous Materials Safety Administration. Pipeline Mileage and Facilities; PHMSA: Washington, DC, 2015; http://www.phmsa.dot.gov/pipeline/library/datastatistics/ pipelinemileagefacilities

(24) Energy Information Administration. Crude Oil Production; EIA: Washington, DC, 2014; http://www.eia.gov/dnav/pet/pet_crd_crpdn_adc_mbblpd_a.htm

CHAPTER 3

Aerial surveys of elevated hydrocarbon emissions from oil and gas production sites

Abstract

Oil and gas (O&G) well pads with high hydrocarbon emission rates may disproportionally contribute to total methane and volatile organic compound (VOC) emissions from the production sector. In turn, these emissions may be missing from most bottom-up emission inventories. We performed helicopter-based infrared camera surveys of more than 8,000 O&G well pads in seven U.S. basins to assess the prevalence and distribution of high-emitting hydrocarbon sources (detection threshold ~1–3 g s^{-1}) . The proportion of sites with such high-emitting sources was 4% nationally but ranged from 1% in the Powder River (Wyoming) to 14% in the Bakken (North Dakota). Emissions were observed three times more frequently at sites in the oil-producing Bakken and oil-producing regions of mixed basins (p<0.0001, χ^2 test). However, statistical models using basin and well pad characteristics explained 14% or less of the variance in observed emission patterns, indicating that stochastic processes dominate the occurrence of high emissions at individual sites. Over 90% of almost 500 detected sources were from tank vents and hatches. Although tank emissions may be partially attributable to flash gas, observed frequencies in most basins exceed those expected if emissions were effectively captured and controlled, demonstrating that tank emission control systems commonly underperform. Tanks represent a key mitigation opportunity for reducing methane and VOC emissions.

Introduction

Hydrocarbon emissions from oil and gas (O&G) facilities pose multiple risks to the environment and human health. Methane, the primary constituent of natural gas, is a short-lived greenhouse gas with 28 — 34 and 84 — 86 times the cumulative radiative forcing of carbon dioxide on a mass basis over 100 years and 20 years, respectively.[1] Burning natural gas instead of other fossil fuels may increase net radiative forcing for some time, even if carbon dioxide emissions decline, depending on the loss rate of methane across the O&G supply chain.[2] O&G hydrocarbon emissions also include volatile organic compounds (VOCs), which are defined by the United States Environmental Protection Agency (U.S. EPA) as photochemically reactive organic compounds excluding methane and ethane. VOCs contribute to regional ozone formation and have been linked to elevated ozone levels in several O&G producing regions.[3,4] Certain VOCs such as benzene are toxic and may be connected to increased risk of cancer and respiratory disease in some areas with O&G development.[5,6]

Hydrocarbons (HC) can be emitted from vented, fugitive, or combustion sources. Vented HC emissions are intentional releases of natural gas from blowdowns (releasing gas to depressurize equipment for maintenance or safety) or sources that emit as part of routine operations such as pneumatic controllers. Fugitive HC emissions are unplanned releases from equipment leaks or malfunctioning equipment. Combustion HC emissions include uncombusted hydrocarbons in the exhaust of combustion sources such as compressor engines and flares. HC emissions can also occur from storage tanks for oil, natural gas condensate, and produced water. Tanks can be the source of both vented emissions, such as flashing losses when liquids are dumped from high-pressure separators to atmospheric pressure tanks, and fugitive emissions caused by malfunctioning separators or control devices. Unlike emissions of raw natural gas, which are primarily composed of methane, oil and condensate tank flashing emissions tend to be dominated by heavier alkanes such as propane and butane.[7]

Recent studies have used two broad approaches to estimate methane or VOC emissions: top-down methods that quantify emissions at the regional or larger scale at one or more points in time, and bottom-up methods that use activity data and emission factors to scale up component- or facility-level measurements to generate emission inventories. Generally, top-down estimates of methane emissions

66

have been greater than bottom-up estimates.[8,9] In the Barnett Shale, a coordinated campaign with simultaneous top-down and bottom-up methods was able to reconcile aircraft mass balance estimates of regional O&G methane emissions with a custom emission inventory based on local and national facility-level measurements.[10,11] Compared to traditional bottom-up inventories, the coordinated campaign inventories estimated higher emissions due to more comprehensive activity data and the inclusion of high emission 'super-emitter' sites in the development of emission factors.[11,12]

Many types of O&G facilities have highly skewed emission distributions with a small fraction of sites contributing the majority of emissions.[13–17] These high emission facilities, often referred to as superemitters, may include some sites with persistent emissions and others with intermittent episodes of large releases.[18] High emission rates are likely due to both fugitive emissions caused by malfunctions and vented emissions such as tank flashing or blowdowns. The identification and mitigation of high emission sites is critical to reducing regional emissions since these facilities contribute a large portion of total O&G emissions.[18,19] If the identity of these sites can be predicted, then it would be effective to focus mitigation efforts on sites with characteristics most often associated with high emissions. However, if the occurrence of high emissions is stochastic, then the only viable mitigation solution would be frequent or continuous monitoring of all sites in order to quickly identify and mitigate those with excess emissions.

A common method to detect HC leaks at O&G facilities is optical gas imaging, which has been proposed by U.S. EPA as a regulatory requirement for new and modified sources.[20] Since methane and other HC emissions are invisible to the naked eye, infrared (IR) cameras are used to visualize HC plumes .[21] IR cameras cannot differentiate individual HC species nor quantify emissions under field conditions, but their ability to identify the exact location of an emission source is highly valuable for mitigation. A skilled technician on the ground can use an IR camera to quickly survey thousands of components at an O&G facility for leaks.[22] Helicopter-based IR camera surveys have been used by operators and regulatory agencies to inspect large numbers of sites for high emission rate sources that may indicate equipment issues or noncompliance with environmental regulations.[23]

In this study, we use data collected during helicopter-based IR camera optical gas imaging surveys of more than 8,000 O&G well pads to assess the prevalence and distribution of high-emitting HC sources in

seven U.S. O&G basins. Survey data were analyzed to determine patterns and statistical relationships of observed emissions with well pad and operator parameters. In turn, observed frequencies of high emission sources were compared to predicted frequencies of observable tank flashing emissions with and without controls to assess if detected emission sources indicate the presence of malfunctioning emission control systems.

Methods

Survey areas were selected by stratified random sampling in seven U.S. O&G basins accounting for 33% and 39% of U.S. oil and gas production, respectively: Bakken (North Dakota/Montana), Barnett (north central Texas), Eagle Ford (south Texas), Fayetteville (Arkansas), Marcellus (Appalachian Basin), Powder River (Wyoming/Montana), and Uintah (Utah). Sub-regions in each basin were selected based on their suitability for helicopter surveys (<1,500 meters above sea level, unrestricted airspace) and subdivided into 10 x 10 km grid cells. Due to their large size, sub-regions in the Bakken, Marcellus, and Powder River were centered on areas with active drilling in northwest North Dakota, southwest Pennsylvania, and eastern Wyoming, respectively. Data on well pad characteristics for each of these sub-regions were obtained for wells with an active status from the production database Drillinginfo, which contains data compiled and cleaned from state databases.[24]

One or two defining characteristics were identified for each region that best characterized the heterogeneity of the basin's O&G production and could be the basis for stratified sampling. The selected strata were gas-oil ratio (GOR) in produced fluids (Barnett and Uintah), well age (Bakken), a combination of well age and GOR (Marcellus), and well type of oil, gas, or coal-bed gas (Powder River). These strata were chosen to reflect the distinguishing characteristics in each region (e.g., GOR does not vary greatly in the Bakken, so no meaningful stratification is possible along that dimension). Parameter thresholds separating strata were selected independently for each basin to divide grid cells into two or three quantiles of average parameter values. After assigning grids to strata based on their average parameters, a list of grids in each stratum was randomly selected for survey. In two basins this design was not followed. In the Fayetteville, a single 20 x 20 km area was selected due to limited survey time and homogenous production across the basin (dry gas without oil). In the Eagle Ford, two unstratified 40 x 15

km survey areas that each covered the basin's broad range in GOR were selected to facilitate efficient measurements by additional research aircraft. A map of surveyed basins is shown in Figure S1.

Survey area boundaries were provided to a professional firm with extensive experience performing leak detection surveys of O&G sites from a helicopter using optical gas imaging (Leak Surveys, Inc.).[25] Flights occurred from June to October 2014 using an R44 helicopter. The survey team identified O&G well pads, compressor stations, and small gas processing plants in the survey areas; for this paper, only data from active well pads are included in the analysis. Camera operators used a FLIR GasFindIR infrared camera to visually survey sites for detectable hydrocarbon plumes at an elevation of approximately 50 m above ground level. At each site with detected emissions, the survey team reported the site's latitude/longitude and the number and equipment type of each observed emission source. Additionally, an IR video was recorded at each site with detected emissions, typically by circling the site and focusing on observed emission sources for 20 to 80 seconds. Videos were reviewed by the lead author to verify the number and type of detected sources.

Two independent methods were used to estimate the minimum detection limit of optical gas imaging with an IR camera deployed from the survey helicopter. First, an operator in the Fayetteville performed a controlled release of dry natural gas (97% CH_4) from a pipeline pig receiver at a midstream facility while being observed by the helicopter survey team from a typical survey position during cloudy conditions. A variable orifice was used to release natural gas at three rates. These rates were quantified by the bagging method at ~3, 8, and 27 g s^{-1}, respectively. The helicopter survey team recorded observable plumes from all three controlled release tests with the lowest release rate producing only faint images that appeared to represent the detection threshold under test conditions. Second, an aircraft with a methane analyzer used the atmospheric budget method to quantify methane emissions at 19 well pads and compressor stations within one hour of detection by the helicopter survey team (See SI for methodological details). Measured site emission rates ranged from 1 to 24 g CH_4 s^{-1} with 84% of central estimates above 3 g CH_4 s^{-1} (Table S1). Additionally, the helicopter survey team qualitatively ranked the size of emission sources based on the apparent size and density of plumes, but there was no correlation between the qualitative magnitude of emission sources estimated from experienced camera operators and the quantified methane emission

rates; potential reasons are discussed in the SI. Variability in the IR camera's sensitivity to different hydrocarbons (HC) is expected to impact the detection limit. The GasFindIR camera can detect at least 20 different HCs with differing functionality and has the highest sensitivity to alkanes; the reported minimum detectable emission rate under controlled conditions is 2 – 4 times lower for propane than methane under controlled conditions.[21] While there may be differences in the ratio of minimum detectable emissions rates in the field compared to controlled conditions, a ratio of 3 was chosen as representative of the increased sensitivity of the camera to higher molecular weight HCs. Therefore, the helicopter survey detection limit was assumed to be ~3 g HC s^{-1} for dry gas sources with emissions composed primarily of methane and ~1 g HC s^{-1} for sources such as tanks with emissions composed primarily of higher HCs such as propane. The detection limit of the IR camera is also affected by wind speed. We assessed the average wind speed during surveys based on hourly data during daytime hours from local weather stations. Average wind speed ranged from 2.7 m s^{-1} in the Uintah to 6.4 m s^{-1} in the Powder River (Table S2). Based on the power law relationship between wind speed and detection limit reported in Benson et al., the difference in wind speed would cause the average detection limit to be 3 – 4 times higher in the Powder River compared to the Uintah.[21] Therefore, variability in wind speed contributes uncertainty to the detection limit of a similar magnitude as variable gas composition.

Because survey results were reported for unique well pads rather than by individual well (i.e., many sites had multiple wells), the latitude/longitude of individual wells in surveyed areas were used to aggregate wells into pads by spatially joining all active wells within a 50 meter buffer distance.[18] For each pad, well-level data were used to determine the operator, well production type (oil, gas, oil & gas, coal bed methane), well drill type (vertical, horizontal, directional), number of wells, pad age (months since initial production of newest well), gas production, hydrocarbon liquid production, and water production.[24] Hydrocarbon liquid production includes both crude oil and natural gas condensate; for this analysis, the term "oil" is used to refer to all hydrocarbon liquids. Water production data were not available for individual wells in the Fayetteville or Marcellus basins. Parameters were specific to the same survey month for all basins except the Marcellus, for which only annual and semiannual data were available for conventional and unconventional wells, respectively. In addition to pad-specific parameters, operator-specific parameters were calculated for each basin based on operators' full population of wells in each

basin. Surveyed sites with detected emissions were matched to individual pads in the survey area using the reported latitude/longitude as well as Google Earth imagery.

The helicopter-based team surveyed 8,220 well pads located throughout an area of 6,750 km^2. Average well pad characteristics by basin and strata are summarized in Table S3. The average number of wells per pad ranged from 1.1 in the Uintah to 2.7 in the Fayetteville. Well pads were newest in the Fayetteville (average age of newest well on each pad of 4.1 years) and oldest in the Barnett (13.4 years). Average gas production ranged from 65 Mcf pad^{-1} day^{-1} in the Uintah to 1,438 Mcf pad^{-1} day^{-1} in the Fayetteville. Average oil production ranged from 0 bbl pad^{-1} day^{-1} in the Fayetteville to 312 bbl pad^{-1} day^{-1} in the Eagle Ford. For the basins with oil production, GOR was lowest in the Bakken (1.2 Mcf bbl^{-1}) and highest in the Marcellus (153 Mcf bbl^{-1}). To assess the representativeness of surveyed sites, we compared these parameters between surveyed sites and the total population of active wells in each basin in 2014 (Table S4). For almost all parameters, surveyed sites had statistically different distributions than the entire basin (Kolmogorov-Smirnov $p > 0.05$) but the percent difference for most values was <25% from the basin mean and almost always within 50%. In all basins, surveyed wells were younger than the full population; in the Bakken, Barnett, Eagle Ford, Marcellus, and Powder River, surveyed wells had higher gas production and/or oil production than the basin average. These slight biases likely resulted from selecting sub-regions with active drilling to include young sites in our survey areas. Overall, our sampled strata account for the full range of diversity within and across basins and are appropriate for assessing patterns in high emissions.

Pearson's correlation coefficients (r) were used to assess correlation between the presence (non-detect = 0, detect = 1) or number of detected emissions by source type and pad or operator parameters. Binomial generalized linear models (GLM), also known as logistic regression models, were used to predict the probability of detected emissions at a well pad (P_{detect}) from site and operator parameters. Analysis of variance models and Tukey's Honest Significance Difference test were used to assess significant differences in P_{detect} among basins, strata, well type, and drill type. Poisson GLMs were used to predict the number of detected sources by emission type at each pad. For the full dataset and individual basins, several single parameter and multi-parameter GLMs were evaluated based on their simplicity, Akaike

Information Criteria, Pearson's r, and Hosmer-Lemeshow goodness of fit between observed and predicted values to select models meaningful for explaining the effects of parameters on emissions. An alpha level of 0.05 was used to determine statistical significance in all tests. For statistical analyses, percent energy from oil was used as a surrogate for GOR since it has a discrete range and is more normally distributed; this metric was calculated from oil and gas production using an assumed energy content of 5.8 MMBtu bbl^{-1} for oil and 1.05 MMBtu Mcf^{-1} for natural gas.[26]

Results & Discussion

A total of 494 unique high emissions sources at 327 well pads were detected by the helicopter survey team out of 8,220 surveyed well pads in seven basins. The percentage of total well pads with detected HC emissions (P_{detect}) was 4% but ranged from 1% in the Powder River to 14% in the Bakken (Table 1). There were statistically significantly differences in P_{detect} by basin with the Bakken higher than all other basins (see Table 1 for full pair-wise comparisons). Emissions were more often observed in oil-producing areas with an average P_{detect} of 13% in the Bakken and low gas-to-oil ratio strata of mixed production basins (p<0.0001, χ^2 test). For example, in the Barnett, 21% of well pads in the low GOR strata showed detectable emissions compared to <1% of sites in the high GOR strata (Table 1). There were also significant differences in P_{detect} by well production type (oil & gas > oil > gas > coal bed methane) and well drill type (horizontal > directional and vertical).

Tank hatches and tank vents were the most common source type of detected emissions, comprising 92% of observed sources (Table 1). The remaining 8% of detected emission sources were dehydrators, separators, trucks unloading oil from tanks, and unlit or malfunctioning flares. Detected emission sources represent individual release points with HC emission rates exceeding the survey's estimated detection limit of approximately 3 g s^{-1} for CH_4 or 1 g s^{-1} for heavier HC. Given this detection limit, no emissions were observed from equipment leaks, pneumatic controllers, or chemical injection pumps, consistent with two recent studies that observed a maximum emission rate of 1.5 CH_4 g s^{-1} at over 1,000 such measured sources.[27,28]

There are several factors that may account for differences among basins in P_{detect} including operational

practices, emission control regulations, and the mix of produced hydrocarbons. The effect of weather

conditions on the detection limit may also have impacted the frequency of observed emissions. In

particular, the higher average wind speed in the Powder River may have contributed to the low frequency

of observations.

Statistical analyses

There were statistically significant but relatively weak positive correlations between detection and

numerous well pad parameters: well count, gas production, oil production, water production, and percent

energy from oil (Table 2; r = 0.06 to 0.20). Detection was negatively correlated with well age (r = -0.12),

meaning that newer wells were more likely to have detected emissions. The average P_{detect} by decile of

analyzed pad parameters is shown in Figure 1. One binomial generalized linear model, GLM A4,

predicted detection that was not significantly different than observed (Hosmer-Lemeshow > 0.05); this

multi-parameter model used basin, the numerical pad parameters well count, well age, gas production, oil

production, and percent energy from oil, and the interactions of basin with each numerical parameter, to

explain 14% of the variance in P_{detect} (r^2 = 0.14). Three other multi-parameter GLMs had observed and

predicted detections that were statistically different and explained 3 to 8% of the variance: A1, using basin

only; A2, using numerical parameters only; and A3, using basin and the numerical parameters but not

their interactions (Table S5, r^2 = 0.03, 0.07, & 0.08, respectively). The increase in predictive power

indicates that the effect of well pad numerical parameters on P_{detect} varies by basin. For example, in the

Marcellus, Powder River, Barnett, and Uintah basins, which have a mix of produced hydrocarbons with a

wide range of GOR, there was a significant positive effect of percent energy from oil on predicted P_{detect},

while there was no significant effect of this parameter in the basins with more homogenous production.

The most predictive GLM, A4, only explained 14% of the variance, which indicates that the presence of

high emissions was primarily stochastic or driven by operational characteristics not included in this

analysis. Therefore, statistical models have limited utility for predicting the occurrence of individual high

emission sources. However, the relatively weak, statistically significant correlation of several parameters

with P_{detect} does provide some insights into factors affecting the likelihood of high emissions. To assess

73

the effects of well pad characteristics on detection, we evaluated single parameter GLMs for the full dataset and individual basins (Table S6). For the full dataset, the GLMs with the best fit between observed and predicted detection were based on well age ($r^2 = 0.04$), oil production ($r^2 = 0.03$), and percent energy from oil ($r^2 = 0.03$). The relative strength of the effects of parameters on the likelihood of detection can be assessed by the ratio of GLM predicted P_{detect} at the 97.5[th] and 2.5[th] percentile of parameter distributions. For example, predicted probability of detection for a pad at the 97.5[th] percentile of well count ($P_{detect} = 0.11$; 5 wells per pad) is 3.2 times higher than for a pad at the 2.5[th] percentile ($P_{detect} = 0.03$; 1 well per pad). For individual basins, single parameter GLMs with statistically significant fits had the same directional effects as the full dataset but varied in their relative strength. The best fit GLMs were based on well count in the Bakken and Marcellus, well age in the Powder River and Uintah, and oil production in the Barnett and Eagle Ford. In the Fayetteville, no single parameter GLM had a statistically significant fit. Detailed parameters for GLM A4 and single parameter GLMs are reported in Tables S16 & S17.

Other studies have reported a weak positive correlation between gas production and methane emissions.[13,28,29] In a prior study of the Barnett Shale, the top 7% of well pads by gas production were estimated to contribute 29% of total methane emissions; this was attributed to higher absolute emissions yet lower proportional loss rates of produced gas at high production sites.[18] The positive correlation between oil production and emission detection may be related to a higher frequency of tank flashing with increased oil production. Brantley et al. reported that oil production was negatively correlated with methane emissions as part of a multivariate linear regression model, which the authors attributed to lower methane content relative to heavier HCs in gas from oil producing wells.[13] In this study, the opposite effect would be expected since the IR camera detects all HCs with higher sensitivity to heavier HCs. The positive relationship between the number of wells per pad and P_{detect} may be due to greater complexity and potential emission sources at multi-well pads. The negative effect of well pad age, the parameter with the strongest predictive power, is likely related to the inverse relationship between well age and oil and gas production, although all parameters remain significant in multi-parameter GLMs. Pads with a well in its first two months of production had over five times the frequency of detected emissions than older pads ($p < 0.001$). Due to the steep decline in production rates of unconventional wells with age, equipment and

control devices may be undersized for handling this period of maximum production. Although older sites would be expected to have a greater likelihood of malfunctions caused by equipment wear, young sites may have initial issues caused by poor setup that have yet to be detected and repaired.

Similar statistical analyses were performed for basin-specific operator characteristics — there were several statistically significant but weak correlations between P_{detect} and these parameters (Table 2) with the strongest positive and negative correlations for an operator's regional percent energy from oil (r = 0.17) and regional well count (r = -0.11). Binomial GLMs predicting P_{detect} from operator parameters are described in the SI. Relationships between the number of detected emissions by source type and well pad or operator characteristics were also evaluated (Table 2). For tank vents and hatches, the number of detected sources at a pad was most strongly correlated with oil production (r = 0.24 and 0.19). For non-tank sources, correlations were weaker (r = -0.06 to 0.06). Poisson GLMs predicting the number of detected sources by type from pad parameters are described in the SI.

Potential causes of observed emissions

High-emitting sources detected by the survey team may have been caused by both malfunctions and normal operations. For non-tank sources, IR videos provide evidence that most sources were the result of malfunctions or intentional releases. There were fourteen observations of malfunctioning flares that were unlit or operating with poor combustion efficiency. Emissions were detected from the pressure relief valves of four separators; although these pressure relief valves may have been functioning properly for safety purposes, the overpressurization that triggered their release indicates abnormal operations. Eight emission sources were observed from vents associated with trucks unloading oil from tanks, which may be intentional to relieve pressure of gas that is released as oil is pumped into trucks. Fifteen dehydrators were observed to have HC emissions, primarily from still vents that remove water vapor from the water-saturated glycol solution. Based on pad gas production and HC emission factors, no more than three of these dehydrators would be expected to have still vent emissions close to the 1 g s^{-1} detection limit.[30] Therefore, most observed emissions from dehydrators were likely the result of abnormal operations that allowed excess HC to slip through the vent. In addition to the IR videos of individual sources, the very

weak fit between observed and predicted emissions suggests that non-tank emission sources are strongly driven by stochastic processes such as malfunctions.

Attributing tank vent and hatch emissions to malfunctions or normal operations is more difficult due to the many potential causes of tank emissions. As part of normal operations, uncontrolled tanks emit HCs from working, breathing, and flashing losses. Tank working and breathing losses generally are expected to be less than 1 g HC s^{-1}, but emissions in excess of this rate can occur from tank flashing after a separator dumps liquids into a tank.[13,31] As discussed below, the emission rate and frequency of tank flashing emissions can be predicted based on parameters including oil production.

Another routine cause of tank emissions is when wells are vented to unload liquids accumulated in the wellbore, which also releases gas. Emissions from well unloadings can be very large — the average emission rate of over 100 measured unloading events was 111 g CH_4 s^{-1}.[32] U.S. EPA Greenhouse Gas Reporting Program (GHGRP) data were used to estimate the percentage of wells expected to be venting at any one time in surveyed basins.[33] Assuming the duration of unloading events was one hour, 0.24% and 0.15% of wells in the Fayetteville (Arkoma) and Uintah basins would have been venting due to liquids unloading at any one time, respectively; all other surveyed basins are predicted to have had less than 0.1% of wells venting from unloading. Therefore, liquids unloading events likely could be detected by the helicopter survey but only can explain a small fraction of observed tank sources.

Finally, abnormal emissions can occur if a separator dump valve fails to properly close and allows produced gas to flow through the tank instead of the sales line. These sources can have very large emission rates — theoretically up to a well's entire gas production if the valve is stuck fully open. In 2014, operators reported over 7,000 malfunctioning dump valves to the U.S. EPA GHGRP.[33] Based on the reported number of hydrocarbon tanks, approximately 5% of GHGRP tanks were associated with stuck dump valves. Operators do not report the duration of stuck dump valves, but a median duration of 7 days can be back calculated from other GHGRP data. Consequently, less than 0.1% of tanks are expected to have emissions from stuck dump valves at any one time.

Influence of flashing emissions by basin

76

To determine if flashing could account for the observed P_{detect} of tank emission sources, potential HC emissions from tank flashing were estimated for surveyed well pads. Flash emission rates per unit of liquids production vary by parameters such as separator pressure and API gravity (a measure of HC liquid density). Since these values were not known for individual sites, basin-level data were obtained from the U.S. EPA O&G Emission Estimation Tool 2014 version 1, which includes a compilation of best available data from several sources including state regulatory agencies.[34] The tool provides separate emission factors for produced water, condensate, and crude oil (Table S11). For hydrocarbon liquids, a weighted average emission factor was derived from basin-level oil and condensate production. If tanks at a well pad are manifolded together with a common vent, then flash emissions will occur when any well's separator dumps to the tank battery. Therefore, site-level production was used as a conservatively high estimate of flashing emissions. The temporal variability of flash emissions depends on the frequency and duration of separator dumps and duration of subsequent flash gas venting. Brantley et al. reported that a tank at a Denver-Julesburg well pad producing 29 bbl d^{-1} condensate flashed ten times in twenty minutes; the duration of flash events in the study ranged from 30 – 120 seconds.[31] This indicates that although individual flash events are short-lived, some sites may have near continuous tank flashing emissions due to frequent venting from separator dumps. To estimate the percentage of sites expected to have flash emissions ≥ 1 g HC s^{-1} detection limit at any one time, the frequency and emission rate of flash emissions were calculated using two sets of assumptions: continuous emissions at a constant rate or intermittent emissions at the detection limit. Both these estimates use the same daily average emission rate but serve as lower and upper bounds for the fraction of sites with concurrent emissions at or above the detection limit. The effects of these assumptions were tested with a sensitivity analysis including alternative emission factors and a 3 g HC s^{-1} detection limit (Tables S12-S15). In all basins, the range of predicted frequencies of sites with uncontrolled tank flashing emissions ≥ 1 g HC s^{-1} included or exceeded observed frequencies (Figure 2; Fayetteville was excluded due to lack of reported liquids production). This indicates that tank flashing could explain observed emissions in the absence of tank emission control devices.

There are several state and federal regulations that require some oil and condensate storage tanks to control VOC emissions, including in North Dakota, Pennsylvania, Utah, and Wyoming.[35,36] For example, during the time of the survey, U.S. EPA New Source Performance Standard Subpart OOOO required all

tanks that began construction after April 12, 2013 and had a potential to emit ≥ 6 tons per year VOC to install control devices with at least 95% control effectiveness within 60 days of initial production.[37] Tank emission control devices include flares, combustors (enclosed flares), and vapor recovery units. The improper design, construction, or maintenance of tank control devices can reduce the capture or control efficiency of tank control devices.[38] Combustion devices can fail to ignite or have poor combustion efficiency, which causes HC emissions from the combustor stack. Emissions may not be fully captured if control systems are undersized or if condensed liquids in vent lines restrict the flow of gas, which can lead to tank overpressurization that triggers the release of gas from a pressure relief valve or tank hatch. Additionally, tank hatches that are left open accidentally or improperly sealed can allow some portion of vented flash gas to circumvent control devices. To determine if the frequency of observed tank emissions indicates failure of tank control systems, we estimated the percentage of sites expected to be equipped with tank controls by applying basin-level control data from the U.S. EPA O&G Emission Estimation Tool (Table S11).[34] For every surveyed well pad, potential emissions from oil, condensate, and water flashing were estimated with basin-level emission factors. Well pads were ranked by potential emissions and then controls were assumed to be equipped at a fraction of sites equal to the percentage of tanks with flares reported in the tool (28 – 86%). Emissions were assumed to be controlled at the reported basin-level capture efficiency (100%) and control efficiency (91 – 98%).[34] If these assumptions were true, then no emissions should be observed from hatches or vents of controlled tanks since all emissions are captured by the control device, but emissions could be observed exiting control devices if uncombusted HC in flare exhaust exceeds the detection limit.

In the Barnett, Powder River, Marcellus, and Uintah Basins, the observed frequency of well pads with detected tank emissions exceeded the maximum predicted frequency based on controlled tank flashing emissions, while in the Bakken the observed frequency was lower than expected (Figure 2). U.S. EPA recently issued a compliance alert that reports inspectors frequently observe emissions from tank hatches and pressure relief valves.[38] After an inspection of almost a hundred tanks in Colorado found numerous instances of ineffective control systems caused by design issues such as undersized control devices, an O&G operator entered a consent decree with U.S. EPA and the State of Colorado to evaluate and improve their control systems.[39] In the Bakken and Barnett, we inspected Google Earth imagery to assess

the presence of tank control devices at well pads with observed tank emissions — 86% and 56% of well

pads with extant imagery, respectively, had apparent control devices. This study's observation that tank

hatches and vents were the origin of the majority of detected large emission sources, even at controlled

sites, suggests that the U.S. EPA O&G Emission Estimation Tool's assumption of 100% capture

efficiency is inaccurate and incomplete capture of emissions by tank control systems is a widespread

issue.

Policy implications

There are several strategies for reducing emissions from tanks, such as installing vapor recovery towers

or stabilizers to reduce the vapor pressure of liquids entering tanks, properly sizing control equipment,

and maintaining pressure relief valves and tank hatches to prevent leaks. Since this study found a higher

frequency of detected emissions at sites within the first few months of production, controlling tank

emissions as soon as a site enters production could reduce overall emissions. U.S. EPA New Source

Performance Standard Subpart OOOO allows the installation of control devices to be delayed up to 60

days after startup, despite this being a period of maximum production, especially for unconventional

wells characterized by rapid production decline.[37] The use of properly sized control devices as soon as

production is initiated would address a substantial source of emissions. For example, the average Bakken

site produces oil about twice the rate in the first two months as it does during the rest of the first year of

production.[24] Given the evidence reported in this study that the frequency of observed tank emissions is

greater than what would be expected if control systems were functioning effectively, it is clear that

identifying anomalous emissions through regular or continuous monitoring of hydrocarbon emissions

and/or equipment status, such as leak detection and repair programs, would be an effective strategy to

reduce emissions.

Currently, U.S. EPA estimates total annual emissions from all oil and gas production sources of 3.1 Tg

VOC and 2.9 Tg CH_4 with 0.6 Tg CH_4 yr^{-1} attributed to oil and condensate tanks.[40,41] The qualitative

nature of the IR survey data precludes an accurate estimate of hydrocarbon or methane emissions, but

with knowledge of the detection limit of the technology deployed our observations can be used to

estimate a lower bound for tank emissions. Our observation of more than 450 detected tank sources with

emission rates ≥ 1 g HC s^{-1} represent at least 450 g HC s^{-1} (a more likely estimate is ~1,575 g HC s^{-1} based on the median aircraft quantified well pad emission rate of 3.5 g CH_4 s^{-1}). While these emissions likely include both intermittent and continuous sources, the assumption of a relatively constant emission rate across a large number of sites is robust and yields an emission rate of at least 14.2 Gg HC yr^{-1}. Since our observations were limited to summer/fall and daylight hours, we were not able to assess how annual average prevalence may be affected by seasonal or diurnal trends such as higher tank breathing losses during warmer conditions. The 8,220 surveyed well pads include 1.1%, 3.7%, and 4.5% of U.S. active wells, gas production, and oil production, respectively. There is uncertainty in scaling up emissions from our sample given that the representativeness of surveyed wells to the U.S. national population of O&G wells has not been assessed and there are only weak correlations between the prevalence of high emissions and these parameters. However, scaling up by the best fit parameter, oil production, yields a minimum national HC emission rate of 0.32 Tg yr^{-1} from high emission tank sources. This national emission estimate of tank emissions represents a lower bound for high-emitting tanks and excludes common, lower emission rate sources such as tank working and breathing losses. This study provides evidence that the cause of some observed emissions is anomalous conditions rather than routine, intermittent tank flashing. U.S. EPA may be underestimating emissions from O&G tanks by overestimating control effectiveness and failing to comprehensively include abnormal, high emission sources. It is reasonable to assume that tanks are a major contributor to the gap between top-down and bottom-up estimates of O&G CH_4 emissions reported by several studies, as well as to the fat-tail emissions observed in a previous study of the Barnett that closed the gap.[11]

Even though this study found statistically significant correlations between the presence of detected emissions and several well pad and operator parameters, these relationships were weak and GLM models were able to explain less than 15% of the variance. This low degree of predictability indicates that these large emission sources are primarily stochastic and the frequent and widespread inspection of sites to identify and repair high emission sources is critical to reducing emissions. In addition to helicopter-based IR surveys, continuous site-based and mobile leak detection systems may be valuable for quickly identifying these large sources.[13,14,42,43,44] Tank vents and hatches account for the vast majority of high emission sources detected at well pads across the U.S. Although routine tank flashing may be

responsible for some of these emission sources, there is evidence that substantial emissions are caused by abnormal conditions such as ineffective tank control systems. Installing tank control devices on existing sources combined with maintenance and monitoring to assure control systems are operating effectively would be an important step for reducing emissions of methane and VOCs. Tanks and other high emission sources are an important contributor to total hydrocarbon emissions from oil and gas well pads and offer a promising opportunity to reduce emissions, but further reductions targeting the numerous emission sources that are individually smaller but collectively large will also be necessary to minimize the health and climate impacts of oil and gas production.

Table 1. Infrared camera survey results by basin and strata. For the percentage of pads with detected emissions (P_{detect}), letters indicate statistically significant differences among strata within each basin (a—c) and among basins (w—z) as determined by Analysis of Variance models and Tukey's HSD ($p < 0.05$). For example, within the Barnett, P_{detect} in the Low GOR strata is statistically different than the High GOR and Medium GOR strata; the overall Barnett P_{detect} is statistically different than overall P_{detect} of the Bakken, Marcellus, Powder River, and Uintah.

Basin	Strata	Detected Sources				Well Pads with Detected Sources	
		Number	% Tank Vents	% Tank Hatches	% Other Sources	Number	% of pads
Bakken	Young	109	9%	83%	7%	57	14.9%[a]
	Old	61	10%	85%	5%	37	12.4%[a]
	All surveyed	*170*	*9%*	*84%*	*6%*	*94*	*13.8%[w]*
Barnett	High GOR	10	60%	50%	0%	7	0.7%[a]
	Medium GOR	9	22%	67%	11%	6	1.4%[a]
	Low GOR	60	55%	40%	3%	46	20.6%[b]
	All surveyed	*79*	*52%*	*44%*	*4%*	*59*	*3.5%[y]*
Eagle Ford	East	70	61%	34%	3%	29	11.0%[a]
	West	1	0%	100%	0%	1	0.3%[b]
	All surveyed	*71*	*61%*	*35%*	*3%*	*30*	*5.4%[xy]*
Fayetteville	*All surveyed*	*24*	*17%*	*83%*	*0%*	*13*	*4.4%[xyz]*
Marcellus	High GOR, Younger	17	76%	12%	12%	13	1.4%[a]
	High GOR, Older	0				0	0.0%[b]
	Low GOR	15	13%	87%	0%	11	10.7%[c]
	All surveyed	*32*	*47%*	*47%*	*6%*	*24*	*1.2%[z]*
Powder River	Coal Bed Methane	0				0	0.0%[a]
	Oil/CBM mix	0				0	0.0%[a]
	Oil	18	44%	39%	22%	15	11.2%[b]
	All surveyed	*18*	*44%*	*39%*	*22%*	*15*	*1.0%[z]*
Uintah	High GOR	3	67%	0%	33%	3	2.2%[a]
	Medium GOR	59	75%	5%	20%	52	6.3%[ab]
	Low GOR	38	63%	21%	16%	37	8.8%[b]
	All surveyed	*100*	*70%*	*11%*	*19%*	*92*	*6.6%[x]*
All Basins		**494**	**40%**	**52%**	**8%**	**327**	**4.0%**

Table 2. Correlation of well pad and operator parameters with P_{detect} (the detection of emissions at a site; non-detect = 0, detect = 1) or the number of detected sources by type. Well pad parameters represent the individual site. Operator parameters represent all regional well pads operated by the same company as each surveyed site. Reported values are Pearson correlation coefficients (r) that are significantly different than zero ($p < 0.05$).

Parameters		P_{detect}	Total Sources	Tank Vents	Tank Hatches	Non-Tank Sources
well pad parameters	well count	0.15	0.16	0.15	0.10	
	well age	-0.12	-0.10	-0.08	-0.07	-0.03
	gas production	0.12	0.11	0.15	0.04	
	oil production	0.20	0.28	0.24	0.19	
	water production	0.06	0.06	0.04	0.06	
	% energy from oil	0.19	0.16	0.10	0.12	0.06
operator regional parameters	well count	-0.11	-0.09	-0.06	-0.06	-0.05
	gas production	-0.05	-0.03	-0.03		-0.04
	oil production	0.09	0.10	0.06	0.08	
	water production	-0.06	-0.06	-0.04	-0.03	-0.06
	% energy from oil	0.17	0.14	0.08	0.12	0.06

Figure 1. Percentage of well pads with detected emissions by deciles of well pad parameters: a) Well Count (wells per pad), b) Well Age (months since initial production of newest well), c) Gas Production (Mcf/day), d) Oil Production (bbl/day), e) Water Production (bbl/day), and f) % Energy from Oil. The median values of each decile are displayed on the x-axes.
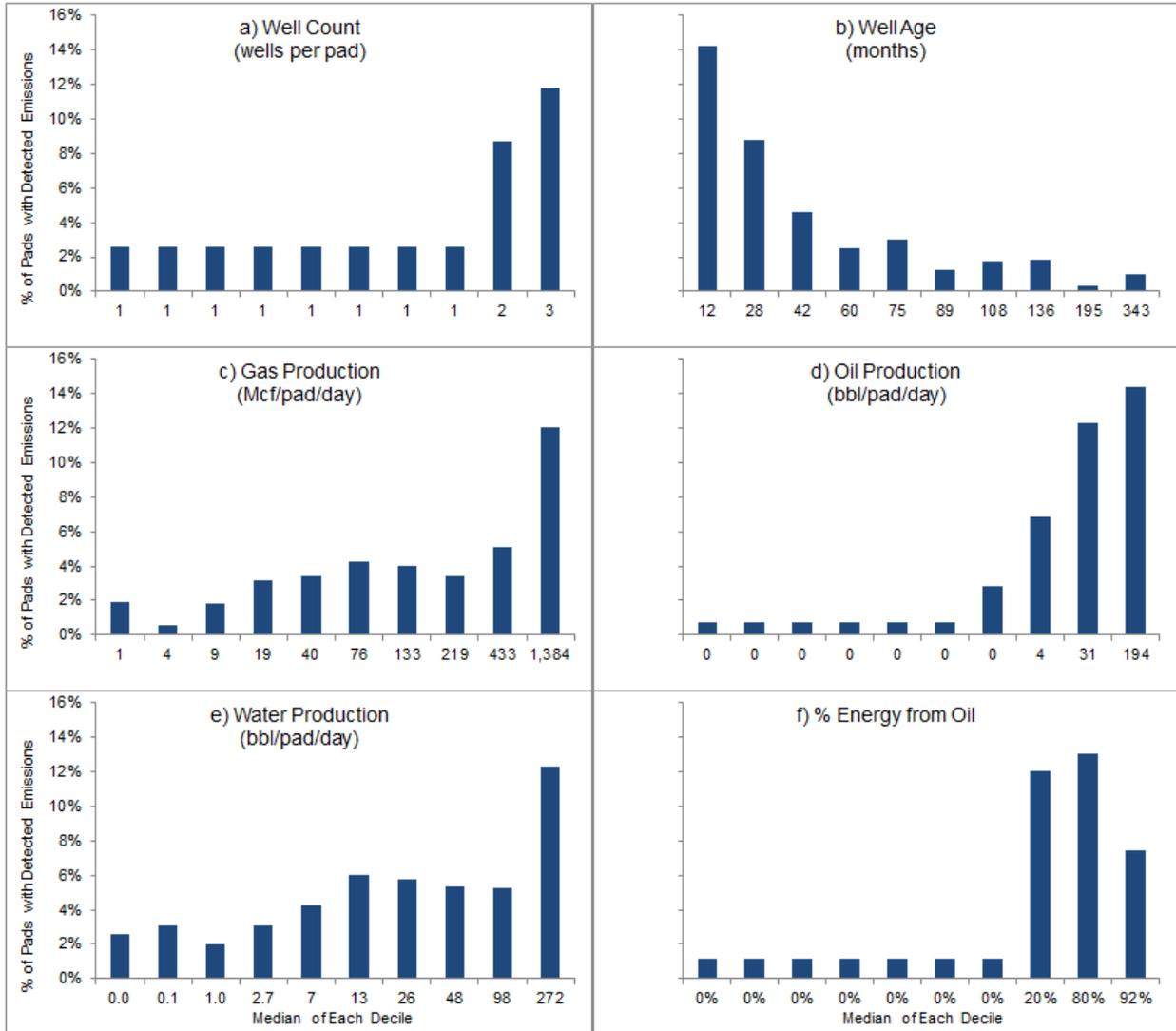
Figure 2. Comparison of the observed and predicted frequencies of well pads with detected tank hydrocarbon emissions assuming an observation threshold of 1 g s^{-1} and basin-level data from the EPA O&G Estimation Tool. Two sets of predicted estimates are provided: red bars reflect predicted frequencies based on potential emissions without controls; green bars reflect the application of controls to the highest emitting tanks (see text for details). Predicted frequencies are shown as a range reflecting different temporal profiles of tank flashing emissions. For several basins and strata, observed frequencies are lower than frequencies predicted without controls but higher than predicted with controls. For example, the combined Uintah observation of 5.8% is within the range predicted for potential emissions but greater than the maximum of 1.5% predicted if all tank control systems were functioning effectively.

Acknowledgements

CHAPTER 3 REFERENCES

1.   Intergovernmental Panel on Climate Change, Stocker TF, editors. Climate change 2013: the physical science basis ; Working Group I contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. New York, NY: Cambridge Univ. Press; 2014. 1535 p.

2.   Alvarez RA, Pacala SW, Winebrake JJ, Chameides WL, Hamburg SP. Greater focus needed on methane leakage from natural gas infrastructure. Proc Natl Acad Sci. 2012 Apr 24;109(17):6435–40.

3.   Kemball-Cook S, Bar-Ilan A, Grant J, Parker L, Jung J, Santamaria W, Mathews J, Yarwood, G. Ozone Impacts of Natural Gas Development in the Haynesville Shale. Environ Sci Technol. 2010 Dec 15;44(24):9357–63.

4.   Carter WPL, Seinfeld JH. Winter ozone formation and VOC incremental reactivities in the Upper Green River Basin of Wyoming. Atmos Environ. 2012 Apr;50:255–66.

5.   Colborn T, Kwiatkowski C, Schultz K, Bachran M. Natural Gas Operations from a Public Health Perspective. Hum Ecol Risk Assess Int J. 2011 Sep;17(5):1039–56.

6.   McKenzie LM, Witter RZ, Newman LS, Adgate JL. Human health risk assessment of air emissions from development of unconventional natural gas resources. Sci Total Environ. 2012 May;424:79–87.

7.   Hendler A, Nunn J, Lundeen J, McKaskle R. VOC emissions from oil and condensate storage tanks [Internet]. Texas Environmental Research Consortium; 2009. Available from: http://files.harc.edu/Projects/AirQuality/Projects/H051C/H051CFinalReport.pdf

8.   Brandt AR, Heath GA, Kort EA, O'Sullivan F, Petron G, Jordaan SM, Tans P, Wilcox J, Gopstein AM, Arent D, et al. Methane Leaks from North American Natural Gas Systems. Science. 2014 Feb 14;343(6172):733–5.

9.   Miller SM, Wofsy SC, Michalak AM, Kort EA, Andrews AE, Biraud SC, Dlugokencky EJ, Eluszkiewicz J, Fischer ML, Janssens-Maenhout G, et al. Anthropogenic emissions of methane in the United States. Proc Natl Acad Sci. 2013 Dec 10;110(50):20018–22.

10.   Harriss R, Alvarez RA, Lyon D, Zavala-Araiza D, Nelson D, Hamburg SP. Using Multi-Scale Measurements to Improve Methane Emission Estimates from Oil and Gas Operations in the Barnett Shale Region, Texas. Environ Sci Technol. 2015 Jul 7;49(13):7524–6.

11.   Zavala-Araiza D, Lyon DR, Alvarez RA, Davis KJ, Harriss R, Herndon SC, Karion A, Kort EA, Lamb BK, Lan X, et al. Reconciling divergent estimates of oil and gas methane emissions. Proc Natl Acad Sci. 2015 Dec 22;112(51):15597–602.

12.   Lyon DR, Zavala-Araiza D, Alvarez RA, Harriss R, Palacios V, Lan X, Talbot R, Lavoie T, Shepson P, Yacovitch TI, et al. Constructing a Spatially Resolved Methane Emission Inventory for the Barnett Shale Region. Environ Sci Technol. 2015 Jul 7;49(13):8147–57.

13.   Brantley HL, Thoma ED, Squier WC, Guven BB, Lyon D. Assessment of Methane Emissions from Oil and Gas Production Pads using Mobile Measurements. Environ Sci Technol. 2014 Dec 16;48(24):14508–15.

14.   Rella CW, Tsai TR, Botkin CG, Crosson ER, Steele D. Measuring Emissions from Oil and Natural Gas Well Pads Using the Mobile Flux Plane Technique. Environ Sci Technol. 2015 Apr 7;49(7):4742–8.

15.   Mitchell AL, Tkacik DS, Roscioli JR, Herndon SC, Yacovitch TI, Martinez DM, Vaughn TL, Williams LL, Sullivan MR, Floerchinger C, et al. Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results. Environ Sci Technol. 2015 Mar 3;49(5):3219–27.

16.   Subramanian R, Williams LL, Vaughn TL, Zimmerle D, Roscioli JR, Herndon SC, Yacovitch TI, Floerchinger C, Tkacik DS, Mitchell AL, et al. Methane Emissions from Natural Gas Compressor Stations in the Transmission and Storage Sector: Measurements and Comparisons with the EPA Greenhouse Gas Reporting Program Protocol. Environ Sci Technol. 2015 Mar 3;49(5):3252–61.

17.   Lamb BK, Edburg SL, Ferrara TW, Howard T, Harrison MR, Kolb CE, Townsend-Small A, Dyck W, Possolo A, Whetstone JR. Direct Measurements Show Decreasing Methane Emissions from Natural Gas Local Distribution Systems in the United States. Environ Sci Technol. 2015 Apr 21;49(8):5161–9.

18.   Zavala-Araiza D, Lyon D, Alvarez RA, Palacios V, Harriss R, Lan X, Talbot R, Hamburg SP. Toward a Functional Definition of Methane Super-Emitters: Application to Natural Gas Production Sites. Environ Sci Technol. 2015 Jul 7;49(13):8167–74.

19.   Jackson RB, Vengosh A, Carey JW, Davies RJ, Darrah TH, O'Sullivan F, Pétron G. The Environmental Costs and Benefits of Fracking. Annu Rev Environ Resour. 2014 Oct 17;39(1):327–62.

20.   Proposed Rule: Oil and Natural Gas Sector: Emission Standards for New and Modified Sources [Internet]. Sep 18, 2015. Available from: https://www.gpo.gov/fdsys/pkg/FR-2015-09-18/pdf/2015-21023.pdf

21.   Benson R, Madding R, Lucier R, Lyons J, Czerepuszko P. Standoff passive optical leak detection of volatile organic compounds using a cooled InSb based infrared imager. AWMA 99th Annual Meeting Papers, 2006 (Vol. 131).

22.   Robinson DR, Luke-Boone R, Aggarwal V, Harris B, Anderson E, Ranum D, Kulp TJ, Armstrong K, Sommers R, McRae TG, et al. Refinery evaluation of optical imaging to locate fugitive emissions. Journal of the Air & Waste Management Association. 2007 Jul 1;57(7):803-10.

23.   Biennial Report to the 84th Legislature, Chapter 1: Agency Highlights, FY2013-2014 [Internet]. Austin, TX: Texas Commision on Environmental Quality; 2015. Available from: http://www.tceq.state.tx.us/publications/sfr/index84/chapter1

24.   Drillinginfo. DI Desktop [Internet]. Austin, TX; 2015. Available from: http://www.didesktop.com/

25.   Leak Surveys, Inc. [Internet]. Available from: http://www.leaksurveysinc.com/

26.   Monthly Energy Review [Internet]. Washington, DC: United States Energy Information Administration; 2015 Sep. Available from: http://www.eia.gov/totalenergy/data/monthly/archive/00351509.pdf

27.   Allen DT, Pacsi AP, Sullivan DW, Zavala-Araiza D, Harrison M, Keen K, Fraser MP, Daniel Hill A, Sawyer RF, Seinfeld JH. Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Pneumatic Controllers. Environ Sci Technol. 2015 Jan 6;49(1):633–40.

28.   Allen DT, Torres VM, Thomas J, Sullivan DW, Harrison M, Hendler A, Herndon SC, Kolb CE, Fraser MP, Hill AD, et al. Measurements of methane emissions at natural gas production sites in the United States. Proc Natl Acad Sci. 2013;110(44):17768–73.

29.   Omara M, Sullivan MR, Li X, Subramanian R, Robinson AL, Presto AA. Methane Emissions from Conventional and Unconventional Natural Gas Production Sites in the Marcellus Shale Basin. Environ Sci Technol. 2016 Jan 29;50(4):2099-107.

30.   Methane Emissions from the Natural Gas Industry. Volume 14: Glycol Dehydrators [Internet]. Gas Research Institute / United States Environmental Protection Agency; 1996 Jun. Available from: http://www3.epa.gov/gasstar/documents/emissions_report/14_glycol.pdf

31.   Brantley HL, Thoma ED, Eisele AP. Assessment of volatile organic compound and hazardous air pollutant emissions from oil and natural gas well pads using mobile remote and on-site direct measurements. J Air Waste Manag Assoc. 2015 Sep 2;65(9):1072–82.

32.   Allen DT, Sullivan DW, Zavala-Araiza D, Pacsi AP, Harrison M, Keen K, Fraser MP, Daniel Hill A, Lamb BK, Sawyer RF, et al. Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Liquid Unloadings. Environ Sci Technol. 2015 Jan 6;49(1):641–8.

33.   Greenhouse Gas Reporting Program [Internet]. Washington, DC: United States Environmental Protection Agency; 2015. Available from: http://www.epa.gov/ghgreporting

34.   Oil and Gas Emission Estimation Tool 2014 Version 1 [Internet]. Washington, DC: United States Environmental Protection Agency; 2015. Available from: ftp://ftp.epa.gov/EmisInventory/2011nei/doc/Tool_and_Report112614.zip

35.   Utah Department of Environmental Quality. Utah General Approval Order for Crude Oil and Natural Gas Well Site and/or Tank Battery, Section II.B.2. [Internet]. Jun 5, 2014. Available from: http://www.deq.utah.gov/Permits/GAOs/docs/2014/6June/DAQE-AN149250001-14.pdf

36.   North Dakota Department of Health. Bakken Pool, Oil and Gas Production Facilities, Air Pollution Control, Permitting and Compliance Guidance [Internet]. May 2, 2011. Available from: http://www.ndhealth.gov/AQ/Policy/20110502Oil Gas Permitting Guidance.pdf

37.   Standards of Performance for Crude Oil and Natural Gas Production, Transmission and Distribution [Internet]. Available from: http://www.ecfr.gov/cgi-bin/text-idx?node=sp40.7.60.oooo

38.   EPA Observes Air Emissions from Controlled Storage Vessels at Onshore Oil and Natural Gas Production Facilities [Internet]. Washington, DC: United States Environmental Protection Agency; 2015. Available from: http://www.epa.gov/sites/production/files/2015-09/documents/oilgascompliancealert.pdf

39.   Consent Decree: United States of America, and the State of Colorado v. Noble Energy, Inc. [Internet]. Denver, CO: United States District Court for the District of Colorado; 2015. Available from: ; http://www.epa.gov/sites/production/files/2015-04/documents/noble-cd.pdf

40.   2011 National Emissions Inventory [Internet]. Washington, DC: United States Environmental Protection Agency; 2013. Available from: http://www3.epa.gov/ttnchie1/net/2011inventory.html

41.   Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2013 [Internet]. Washington, DC: United States Environmental Protection Agency; 2015. Available from: http://www.epa.gov/climatechange/ghgemissions/usinventoryreport.html

42.   Albertson JD, Harvey T, Foderaro G, Zhu P, Zhou X, Ferrari S, Amin MS, Modrak M, Brantley H, Thoma ED. A Mobile Sensing Approach for Regional Surveillance of Fugitive Methane Emissions in Oil and Gas Production. Environ Sci Technol. 2016 Feb 9;50(5):2487-97.

43.   Advanced Research Projects Agency – Energy, United States Department of Energy. Methane Observation Networks with Innovative Technology to Obtain Reductions – MONITOR. 2014; http://arpa-e.energy.gov/sites/default/files/documents/files/MONITOR_ProgramOverview.pdf

44.   Environmental Defense Fund. Methane Detectors Challenge. 2015; https://www.edf.org/energy/natural-gas-policy/methane-detectors-challenge

CHAPTER 3 APPENDIX

Contents:

1. Supporting text, 17 tables, and 2 figures

2. Excel file with calculations used in tank flashing analysis (Tank_flashing_analysis.xlsx)

3. List of surveyed sites by latitude/longitude (surveyed_well_pads.xlsx)

4. Site-level parameter data for well pads in the surveyed areas and basins (Lyon_et_al_data.xlsx)

5. ZIP file with 8 infrared videos and description of observed sources (Example_IR_videos)

Aircraft Quantification of CH$_4$ Emissions

Methane emissions at five Eagle Ford and fourteen Bakken well pads and compressor stations were quantified with the aircraft-based atmospheric budget method at sites within one hour of emission detection by the helicopter survey team. A Mooney TLS fixed-wing aircraft equipped with a Picarro Cavity Ring-Down Spectrometer methane analyzer was used to measure horizontal and vertical gradients in methane concentration around the target sites. The maximum vertical transport of emissions was determined by flying progressively higher until the concentration gradient was no longer observed. Wind speed and direction were measured in real time from the aircraft with horizontal gradients estimated using a least squares linear optimization.[1] Methane emission rates and uncertainty were quantified with the atmospheric budget method (Table S1).[2]

Qualitative Ranking of Detected Emissions

The helicopter survey team reported the apparent magnitude of detected hydrocarbon (HC) emission sources based on a subjective evaluation of the visually observed plume size, density, and velocity on the infrared camera (qualitative ranking of small, medium, large). Methane emission rates estimated with the aircraft-based atmospheric budget method did not correlate with the number of sources or qualitative size categorization by the helicopter survey team (small, medium, and large sites were weighted 1, 2, and 3,

respectively for correlation analysis, r = 0.01). Therefore, the qualitative rankings were excluded from the statistical analyses. Sites that had emission sources qualitatively classified as large but relatively low quantified CH_4 emission rates may have had plumes too close to the ground for the aircraft to fully capture for the atmospheric budget method, or may have been composed primarily of non-methane HCs (which the aircraft did not quantify). The inability to classify infrared videos into meaningful qualitative emission rate categories also may be due to variable effects of wind dispersion on plume size or temporal variation in emission rates between the time of helicopter surveys and aircraft-based quantification.[3,4]

Statistical Analyses of Operator Characteristics

For basin-specific operator characteristics, detection was statistically significantly correlated with a well pad's operator's regional well count, gas production, oil production, water production, and percent energy from oil, but these correlations were weaker than those between P_{detect} and well pad parameters (Table 2). The strongest negative correlation was with operator regional well count (r = -0.11) and strongest positive correlation was with the operator percent energy from oil (r = 0.17). Two binomial GLMs based on operator parameters predicted detection that was not statistically significant different than observed detection (Table S5). GLM B1, a single parameter model based on operator as a categorical parameter, had a similar fit as GLM B5, a multi-parameter model based on basin, operator numerical parameters, and the interaction of basin with numerical parameters (r^2 = 0.10). The single parameter binomial GLMs with the best fit between observed and predicted detection were models based on an operator's regional percent energy from oil (r^2 = 0.03), which had a positive relationship with detection, and regional well count (r^2 = 0.01), which had an inverse relationship (Table S6).

Operators with less than 100 pads in a region on average had detected emissions at 7% of sites compared to 4% for larger operators (p = 0.008). Smaller operators may have had more frequent large emission sources because they have less capital to invest in equipment, staff, or practices for mitigating emissions, such as leak detection and repair programs. The effect of an operator's percent energy from oil may also be related to operational practices – if an operator has a greater focus on oil production, then they may be less incentivized to implement mitigation practices that capture more gas for sale. Different operational practices likely explain the over 30 times greater frequency of detected emissions in the Eagle

Ford eastern survey area compared to the western area, which were similar in average pad parameters. In the western area, half the well pads were operated by a company that typically pipes unseparated oil and gas to central gathering facilities.[5] In contrast, most pads in the eastern area were operated by companies that typically separate and store oil at the pad. The survey team observed tank emissions at four gathering facilities operated by the main company in the western area, which indicates that use of offsite tanks may only move some emissions to downstream locations, but it is likely that this operational practice leads to an overall decrease in supply chain emissions since highly effective controls such as vapor recovery towers would be more cost-effective at large centralized facilities.

Statistical Analyses of Number of Detected Emissions by Source Type

The number of detected emission sources per pad by source type had the same directional correlations as P_{detect} with pad and operator parameters (Table 2). Although many correlations were statistically significant, none were strong (r ≤ 0.28), again demonstrating the dominance of random processes. The number of detected sources from both tank vents and tank hatches was most strongly correlated with pad oil production (r = 0.24 and 0.19, respectively). Compared to tank sources, non-tank emission sources had almost no relationship with well pad parameters – the strongest correlation was with percent energy from oil (r = 0.06). Several multi-parameter Poisson GLMs based on basin and the well pad numerical parameters well age, well count, gas production, oil production, and percent energy from oil predicted a number of observed sources that was not significantly different than observed (Table S7). GLM C1, based on all numerical parameters, was significant for non-tank sources (r^2 = 0.01). GLM C2, based on basin and numerical parameters, was significant for tank vents (r^2 = 0.03). GLM C3, based on basin, numerical parameters, and the interaction of basin with each numerical parameter was significant for all source types with the strongest correlation for total sources (r^2 = 0.07). Single parameter Poisson GLMs predicting the number of sources by type based on well pad parameters were used to evaluate the effects of parameters (Table S8). The best fit between observed and predicted values were based on models with well age or oil production for tank vents (r^2 = 0.02), oil production for tank hatches (r^2 = 0.02), and percent energy from oil for non-tank sources (r^2 = 0.002).

Table S1. Site methane fluxes estimated by the aircraft atmospheric budget method at well pads and compressor stations within one hour of emission source detection by the helicopter-based IR camera survey team and the number and qualitative description of the magnitude of these sources by the survey team.

Basin	Video ID	Site Type	Aircraft CH$_4$ flux estimate (g/s)	Helicopter IR detected sources
Eagle Ford	eagle88	compressor station	24±16	1 small
Bakken	bak04	compressor station	23±3	2 medium, 4 small
Bakken	bak21	compressor station	20±5	6 big
Bakken	bak28	well pad	18±4	1 medium
Bakken	bak45	compressor station	17±2	1 big
Eagle Ford	eagle82	compressor station	14±8	1 big, 2 small
Eagle Ford	eagle46	compressor station	14±34	3 big
Eagle Ford	eagle39	compressor station	13±20	4 medium
Bakken	bak64	well pad	9±3	8 big
Bakken	bak08	compressor station	5±2	2 big
Bakken	bak49	compressor station	5±2	2 big
Eagle Ford	eagle58	compressor station	5±3	1 big
Bakken	bak51	compressor station	4±1	4 big
Bakken	bak58	well pad	4±1	4 medium
Bakken	bak25	well pad	3±1	4 big
Bakken	bak56	well pad	3±1	7 big
Bakken	bak55	compressor station	2±1	2 big
Bakken	bak53	compressor station	2±1	2 medium
Bakken	bak59	well pad	1±1	1 big

Table S2. Average hourly meteorological conditions from 8:00 AM to 6:00 PM during survey days at local weather stations (< 100 km from surveyed wells). Data were obtained from the Iowa Environmental Mesonet archive of METAR airport weather observations.[6]

Basin	Weather Station	Temp (^{0}C)	Wind Speed (m s^{-1})	% hours by cloud cover class				
				Clear	Few	Scattered	Broken	Overcast
Uintah	VEL	26.3	2.7	72%	17%	6%	4%	1%
Bakken	ISN	15.2	3.9	60%	9%	7%	9%	15%
Fayetteville	LRF	26.7	3.1	26%	38%	15%	15%	5%
Eagle Ford	SKF	27.6	4.8	7%	35%	31%	25%	3%
Barnett	DFW	24.7	4.5	2%	67%	17%	15%	0%
Powder River	CPR	22.6	6.4	44%	25%	17%	10%	4%
Marcellus	PIT	21.4	3.9	3%	45%	41%	10%	2%

Table S3. Average characteristics of well pads in surveyed areas by basin and strata

Basin	Strata	Surveyed Area (km^2)	Surveyed Pads	Wells per Pad	Well Age (yrs)	Gas Production (Mcf pad^{-1} day^{-1})	Oil Production (bbl pad^{-1} day^{-1})	Water Production (bbl pad^{-1} day^{-1})	Area GOR (Mcf bbl^{-1})
Bakken	Young	600	383	1.5	3	296	246	132	1.2
	Old	600	299	1.4	7	313	295	163	1.1
	All surveyed	*1,200*	*682*	1.5	*5*	*303*	*267*	*145*	*1.1*
Barnett	High GOR	300	1,028	1.6	11	390	0	33	13,352
	Medium GOR	300	444	1.3	22	197	2	38	125
	Low GOR	300	223	1.9	8	441	16	266	27
	All surveyed	*900*	*1,695*	1.5	*13*	*346*	*3*	*65*	*136*
Eagle Ford	East	600	264	2.2	5	1,014	411	144	2.5
	West	600	287	2.0	6	553	238	117	2.3
	All surveyed	*1,200*	*551*	2.1	*5*	*774*	*321*	*130*	*2.4*
Fayetteville	*All surveyed*	*400*	*295*	2.7	*4*	*1,438*	*0*	*NA*	*NA*
Marcellus	High GOR, Younger Age	800	920	1.3	9	936	0	NA	4,866
	High GOR, Older Age	300	1,042	1.0	14	34	0		14,358
	Low GOR	400	103	3.4	6	2,719	73		37
	All surveyed	*1,500*	*2,065*	1.3	*11*	*570*	*4*		*153*
Powder River	Coal Bed Methane	300	708	1.0	5	202	0	99	4,768
	Oil/CBM mix	300	701	1.0	8	142	3	51	49
	Oil	300	134	1.1	13	112	79	67	1
	All surveyed	*900*	*1,543*	1.1	*7*	*167*	*8*	*74*	*20*
Uintah	High GOR	100	138	1.1	11	157	2	11	82
	Medium GOR	300	831	1.1	13	69	3	24	25
	Low GOR	250	420	1.2	4	28	18	16	2
	All surveyed	*650*	*1,389*	1.1	*10*	*65*	*7*	*20*	*9*
All Basins		**6,750**	**8,220**	**1.4**	**9**	**385**	**48**	**72**	**8**

Table S4. A comparison of average well pad parameters of surveyed sites to the total population of each basin. Asterisks indicate parameters for which the distributions of surveyed and basin sites are not statistically different (Kolmogorov-Smirnov p > 0.05). The basin populations include all active wells with 2014 production. For the Marcellus, the basin population was limited to Appalachian basin wells in Pennsylvania.

Basin	Strata	Wells per Pad	Well Age (yrs)	Gas Production (Mcf pad^{-1} day^{-1})	Oil Production (bbl pad^{-1} day^{-1})	Water Production (bbl pad^{-1} day^{-1})	% Energy from Oil
Bakken	Young	1.5	3	296	246	132	82%
	Old	1.4	7	313	295	163	83%
	All surveyed	1.5	5	303	267	145	83%
	Basin	*1.4**	*10*	*260*	*225*	*245*	*83%*
Barnett	High GOR	1.6	11	390	0.03	33	0%
	Medium GOR	1.3	22	197	2	38	2%
	Low GOR	1.9	8	441	16	266	35%
	All surveyed	1.5	13	346	3	65	5%
	Basin	*1.9*	*19*	*273*	*3*	*114*	*22%*
Eagle Ford	East	2.2	5	1,014	411	144	63%
	West	2	6	553	238	117	63%
	All surveyed	2.1	5	774	321	130	63%
	Basin	*1.9*	*8*	*857*	*222*	*211*	*55%*
Fayetteville	All surveyed	2.7	4	1,438	0	NA	0%
	Basin	*2.3*	*5*	*1,409*	*0*		*0%*
Marcellus	High GOR, Younger Age	1.3	9	936	0.2	NA	2%
	High GOR, Older Age	1.0	14	34	0		0%
	Low GOR	3.4	6	2,719	73		12%
	All surveyed	1.3	11	570	4		2%
	Basin (PA only)	*1.1*	*15*	*251*	*0.4*		*12%*
Powder River	Coal Bed Methane	1.0	5	202	0.04	99	1%
	Oil/CBM mix	1.0	8	142	3	51	7%
	Oil	1.1	13	112	79	67	53%
	All surveyed	1.1	7	167	8	74	8%
	Basin	*1.1**	*15*	*87*	*16*	*97*	*34%*
Uintah	High GOR	1.1	11	157	2	11	4%
	Medium GOR	1.1	13	69	3	24	21%
	Low GOR	1.2	4	28	18	16	79%
	All surveyed	1.1	10	65	7	20	37%
	Basin	*1.3*	*10*	*166*	*14*	*47*	*31%*

Table S5. A comparison of binomial generalized linear models predicting the detection of emissions for the full dataset from basin and numerical pad parameters. For GLM A4, there is no significant difference between observed and predicted values (Hosmer-Lemeshow p > 0.05).

Model	Parameters	AIC	GLM fit (r)	Hosmer-Lemeshow (p)
A1	basin	2520	0.18	2.2E-16
A2	well age + well count + gas production + oil production + % energy from oil	2344	0.26	3.8E-09
A3	basin + well age + well count + gas production + oil production + % energy from oil	2260	0.29	6.4E-04
A4	basin + well age + well count + gas production + oil production + % energy from oil + basin*well age + basin*well count + basin*gas production + basin*oil production + basin*% energy from oil	2105	0.37	0.21

Table S6. Effects of well pad parameters on emission detection probability for the full dataset and individual basins. For each basin, the top value is the Pearson correlation coefficient (r) between observed and fitted values based on single parameter binomial GLMs. Asterisks indicate that a GLM is statistically significant (p < 0.05). The bottom value is the ratio of predicted detection probability at the 97.5^{th} and 2.5^{th} percentiles of the parameters value in each basin. For parameters with negative effects, the inverse ratios are shown and indicated by parentheses.

	Well Count	Well Age	Gas Production	Oil Production	Water Production	% Energy from Oil
Bakken	0.22*	0.11*	0.21*	0.19*	0.08*	0.03
	2.4	(6.9)	2.6	2.6	1.6	(1.4)
Barnett	0.05*	0.26*	0.12*	0.45*	0.01	0.17*
	2.4	(4.4E+3)	3	5.9	1.1	7.1
Eagle Ford	0.1*	0.12*	0.31*	0.36*	0.12*	0.00
	2.7	(1.1E+5)	4.5	4.9	2.7	(1.1)
Fayetteville	0.07	0.02	0.00	NA	NA	NA
	2.8	(1.4)	1.0			
Marcellus	0.38*	0.31*	0.27*	0.02*	NA	0.05*
	12	(6.2E+7)	1.9	1.3		1.7
Powder River	0.03*	0.26*	0.05*	0.07*	0.01	0.19*
	2.8	(2.6E+11)	3.4	1.1	1.5	45
Uintah	0.02	0.15*	0.04	0.12*	0.02	0.08*
	1.2	(61)	1.4	2.6	1.3	2.1
All Basins	**0.12***	**0.20***	**0.09***	**0.17***	**0.02***	**0.17***
	3.2	**(1.1E+3)**	**1.3**	**1.9**	**1.1**	**6.0**

Table S7. A comparison of binomial generalized linear models predicting the detection of emissions for the full dataset from operator or numerical operator parameters. For GLMs B1 and B5, there are no significant differences between observed and predicted values (Hosmer-Lemeshow p > 0.05).

Model	Parameters	AIC	GLM fit (r)	Hosmer-Lemeshow (p)
B1	operator (categorical)	2578	0.31	1
B2	basin + operator (categorical)	2554	0.31	5.9E-03
B3	operator well count + operator gas production + operator oil production + operator % energy from oil	2526	0.18	3.6E-11
B4	basin + operator well count + operator gas production + operator oil production + operator % energy from oil	2460	0.20	9.2E-07
B5	basin + operator well count + operator gas production + operator oil production + operator % energy from oil + basin interactions with numerical parameters	2238	0.31	0.28

Table S8. Effects of operator regional parameters on emission detection probability at their well pads for the full dataset and individual basins. For each basin, the top value is the Pearson correlation coefficient (r) between observed and fitted values based on single parameter binomial generalized linear models. Asterisks indicate that a GLM is statistically significant. The bottom value is the ratio of predicted detection probability at the 97.5th and 2.5th percentiles of the parameters value in each basin. For parameters with negative effects, the inverse ratios are shown and indicated by parentheses.

	Operator Well Count	Operator Gas Production	Operator Oil Production	Operator Water Production	Operator % Energy from Oil
Bakken	0.15*	0.13*	0.14*	0.15*	0.00
	(3.5)	(3.0)	(3.5)	(3.7)	(1.0)
Barnett	0.13*	0.13*	0.29*	0.24*	0.11*
	(8.4)	(8.2)	11	9.1	2.6
Eagle Ford	0.02	0.07	0.04	0.01	0.01
	(1.3)	2.4	2.0	(1.2)	(1.3)
Fayetteville	0.08	0.08	NA	NA	NA
	2.7	2.7			
Marcellus	0.09*	0.11*	0.08*	NA	0.04
	12	11	3.9		3.0
Power River	0.13*	0.13*	0.13*	0.14*	0.18*
	(86)	(280)	25	(470)	101
Uintah	0.06*	0.09*	0.03	0.08*	0.05*
	(3.1)	(8.3)	1.3	(4.6)	1.6
All Basins	**0.11***	**0.05***	**0.07***	**0.09***	**0.16***
	(8.1)	**(2.3)**	**2.9**	**(3.7)**	**5.4**

Table S9. Comparison of Poisson generalized linear models predicting the number of detected emission sources by type for the full dataset. There are no significant differences between observed and predicted values for total sources based on GLM C3, for tank vents based on GLMs C2 and C1, for tank hatches based on GLM C3, and forother sources based on GLMs C1 and C3 (Hosmer-Lemeshow p > 0.05).

Model	Parameters	Source type	AIC	GLM fit (r)	Hosmer-Lemeshow (p)
C1	well count + well age + gas production + oil production + % energy from oil + operator pad count	total sources	3350	0.23	2.2E-13
		tank vents	1645	0.21	1.1E-7
		tank hatches	2056	0.22	2.2E-7
		other sources	492	0.15	0.11
C2	basin + well age + well count + gas production + oil production + % energy from oil + operator pad count	total sources	3253	0.26	1.0E-6
		tank vents	1557	0.18	0.05
		tank hatches	1893	0.24	0.04
		other sources	483	0.15	0.03
C3	basin interactions + basin + well age + well count + gas production + oil production + % energy from oil + operator pad count	total sources	3026	0.27	0.15
		tank vents	1484	0.24	0.12
		tank hatches	1761	0.21	0.64
		other sources	491	0.15	0.46

Table S10. Effects of well pad parameters on predicted number of detected emissions by source type for the full dataset. For each source type, the top value is the Pearson correlation coefficient (r) between observed and fitted values based on single parameter Poisson generalized linear models. Asterisks indicate that a GLM is statistically significant. The bottom value is the ratio of predicted number of detected sources at the 97.5[th] and 2.5[th] percentiles of the parameters values. For parameters with negative effects, the inverse ratios are shown and indicated by parentheses.

All Basins	well count	well age	gas production	oil production	water production	% energy from oil
total sources	0.08*	0.19*	0.04*	0.14*	0.01*	0.14*
	3.7	(9.2E+3)	1.2	1.4	1.1	8.6
tank vents	0.08*	0.15*	0.07*	0.15*	0.01*	0.09*
	4.0	(2.9E+3)	1.2	1.4	1.1	6.0
tank hatches	0.04*	0.15*	0.01*	0.09*	0.01*	0.10*
	3.7	(1.1E+5)	1.2	1.4	1.1	12
other sources	0.00	0.04*	0.01	0.00	0.02	0.05*
	1.1	(31)	1.1	(1.0)	(5.8)	6.4

Table S11. Tank emission factors, prevalence of controls, capture efficiency, and control efficiency for surveyed basins used in tank flashing analysis. Data are based on the EPA O&G Emission Estimation Tool 2014 Version 1.[7]

Basin	Flashing Emission Factor (g HC bbl^{-1})				% HC liquids tanks with flare	% capture efficiency	% control efficiency
	Produced Water	Oil	Condensate	HC liquids (weighted average)			
Fayetteville	53	1,545	4,666	1,617	NA	100%	98%
Bakken	64	3,459	12,894	3,460	83%	100%	91%
Powder River	60	916	277,381	999	86%	100%	97%
Uintah	60	1,652	7,909	1,800	37%	100%	98%
Marcellus	60	797	8,342	1,263	62%	100%	97%
Eagle Ford	60	795	9,831	817	78%	100%	98%
Barnett	60	795	8,722	959	28%	100%	98%

Table S12. Sensitivity analysis of predicted percentage of well pads with potential hydrocarbon (HC) emissions above 1 g HC s^{-1}. Emissions are calculated separately for produced water, HC liquids, and the combination of water and HCs using basin-level emission factors (EF) from the EPA O&G Estimation Tool 2014 version 1. Three different emission factors (EF) are used HC liquids: oil, condensate, and a weighted EF based on basin-level oil and condensate production. For each emission estimate, the % of sites with emissions above the threshold is calculated based on two temporal profiles: continuous emissions at a constant emission rate (con.) and intermittent emissions at the threshold rate (int.).

Basin	Strata	Observed	water		HC liquids (weighted HC EF)		HC liquids (oil EF)		HC liquids (condensate EF)		total liquids (weighted HC EF)	
			con.	int.	con.	int.	con.	int.	con.	int.	con.	int.
Bakken	Young	11.4%	2.7%	10.4%	87.0%	95.4%	86.0%	94.1%	97.7%	98.4%	87.0%	95.4%
	Old	12.8%	1.0%	8.9%	95.6%	98.2%	95.6%	98.1%	98.7%	99.4%	95.6%	98.2%
	Combined	12.2%	1.8%	9.5%	91.8%	97.0%	91.3%	96.4%	98.2%	98.9%	91.8%	97.0%
Barnett	High GOR	0.7%	0.1%	1.3%	0.1%	1.4%	0.0%	0.0%	0.0%	0.3%	0.1%	1.4%
	Medium GOR	1.4%	0.9%	2.3%	1.1%	3.5%	0.2%	1.4%	2.3%	8.4%	1.1%	3.5%
	Low GOR	19.3%	2.7%	9.8%	7.6%	24.4%	1.8%	13.3%	31.8%	50.9%	7.6%	24.4%
	Combined	3.3%	0.6%	2.7%	1.4%	4.9%	0.3%	2.1%	4.8%	9.1%	1.4%	4.9%
Eagle Ford	East	10.6%	1.9%	9.1%	56.4%	72.6%	55.7%	71.4%	86.7%	89.1%	56.4%	72.6%
	West	0.3%	0.7%	7.9%	43.6%	62.8%	42.9%	60.6%	81.2%	85.2%	43.6%	62.8%
	Combined	5.3%	1.3%	8.5%	49.7%	67.5%	49.0%	65.8%	83.8%	87.1%	49.7%	67.5%
Marcellus	Young High GOR	10.7%	0.0%	1.8%	37.9%	51.8%	26.2%	43.9%	64.1%	70.5%	37.9%	51.8%
	Old High GOR	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Low GOR	1.3%	0.0%	0.6%	0.0%	0.9%	0.0%	0.2%	0.4%	0.8%	0.0%	0.9%
	Combined	1.1%	0.0%	0.4%	1.9%	3.0%	1.3%	2.3%	3.4%	3.9%	1.9%	3.0%
Powder River	CBM	0.0%	0.0%	6.8%	0.0%	6.9%	0.0%	0.0%	0.7%	0.7%	0.0%	6.9%
	Oil/CBM	0.0%	0.0%	3.5%	0.7%	5.0%	0.7%	1.6%	8.3%	8.3%	0.7%	5.0%
	Oil	9.0%	0.7%	4.5%	12.7%	24.1%	11.9%	21.4%	67.9%	68.6%	12.7%	24.1%
	Combined	0.8%	0.1%	5.1%	1.4%	7.5%	1.4%	2.6%	10.0%	10.0%	1.4%	7.5%
Uintah	High GOR	1.4%	0.0%	0.7%	0.0%	4.7%	0.0%	3.7%	5.1%	10.5%	0.0%	4.7%
	Medium GOR	4.8%	0.0%	1.7%	0.2%	7.3%	0.1%	5.2%	7.5%	20.3%	0.2%	7.3%
	Low GOR	7.4%	0.0%	1.1%	6.9%	35.6%	4.8%	32.3%	54.3%	75.1%	6.9%	35.6%
	Combined	5.3%	0.0%	1.4%	2.2%	15.6%	1.5%	13.2%	21.4%	35.9%	2.2%	15.6%

Table S13. Sensitivity analysis of predicted percentage of well pads with controlled hydrocarbon (HC) emissions above 1 g HC s^{-1}. Emissions are calculated separately for produced water, HC liquids, and the combination of water and HCs using basin-level emission factors (EF) from the EPA O&G Estimation Tool 2014 version 1. Three different emission factors (EF) are used HC liquids: oil, condensate, and a weighted EF based on basin-level oil and condensate production. Controlled emissions are estimated by applying capture efficiency and control efficiency to a subset of highest emitting tanks based on basin-level control data from the EPA O&G Estimation Tool. For each emission estimate, the % of sites with emissions above the threshold is calculated based on two temporal profiles: continuous emissions at a constant emission rate (con.) and intermittent emissions at the threshold rate (int.).

Basin	Strata	Observed	HC liquids (weighted HC EF)		HC liquids (oil EF)		HC liquids (condensate EF)		total liquids (weighted HC EF)	
			con.	int.	con.	int.	con.	int.	con.	int.
Bakken	Young	11.4%	33.1%	60.9%	33.1%	60.8%	76.3%	92.9%	34.1%	62.4%
	Old	12.8%	29.2%	58.0%	29.2%	58.0%	75.2%	94.3%	29.5%	58.3%
	Combined	12.2%	30.9%	59.2%	30.9%	59.2%	75.7%	93.7%	31.5%	60.1%
Barnett	High GOR	0.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.2%
	Medium GOR	1.4%	0.0%	0.1%	0.0%	0.1%	0.0%	0.7%	0.0%	0.2%
	Low GOR	19.3%	0.0%	0.4%	0.0%	0.3%	0.0%	3.7%	0.0%	0.8%
	Combined	3.3%	0.0%	0.1%	0.0%	0.1%	0.0%	0.7%	0.0%	0.3%
Eagle Ford	East	10.6%	0.4%	8.6%	0.4%	8.4%	28.0%	51.2%	0.4%	8.9%
	West	0.3%	0.3%	5.7%	0.3%	5.5%	23.0%	44.7%	0.3%	6.1%
	Combined	5.3%	0.4%	7.1%	0.4%	6.9%	25.4%	47.8%	0.4%	7.4%
Marcellus	Young High GOR	10.7%	0.0%	3.1%	0.0%	1.9%	2.9%	19.6%	0.0%	3.1%
	Old High GOR	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Low GOR	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
	Combined	1.1%	0.0%	0.2%	0.0%	0.1%	0.1%	1.0%	0.0%	0.2%
Powder River	CBM	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.4%	0.0%	0.2%
	Oil/CBM	0.0%	0.0%	0.1%	0.0%	0.1%	3.6%	5.8%	0.0%	0.2%
	Oil	9.0%	0.0%	2.9%	0.0%	2.6%	33.6%	46.5%	0.0%	3.0%
	Combined	0.8%	0.0%	0.3%	0.0%	0.3%	4.6%	6.9%	0.0%	0.5%
Uintah	High GOR	1.4%	0.0%	0.9%	0.0%	0.8%	0.0%	3.9%	0.0%	1.5%
	Medium GOR	4.8%	0.0%	1.1%	0.0%	1.0%	0.0%	4.8%	0.0%	1.4%
	Low GOR	7.4%	0.0%	1.8%	0.0%	1.7%	0.0%	8.1%	0.0%	1.9%
	Combined	5.3%	0.0%	1.3%	0.0%	1.2%	0.0%	5.7%	0.0%	1.5%

The header row above the sub-columns reads: % of sites with controlled tank emissions > 1 g HC s^{-1}

Table S14. Sensitivity analysis of predicted percentage of well pads with potential hydrocarbon (HC) emissions above 3 g HC s^{-1}. Emissions are calculated separately for produced water, HC liquids, and the combination of water and HCs using basin-level emission factors (EF) from the EPA O&G Estimation Tool 2014 version 1. Three different emission factors (EF) are used HC liquids: oil, condensate, and a weighted EF based on basin-level oil and condensate production. For each emission estimate, the % of sites with emissions above the threshold is calculated based on two temporal profiles: continuous emissions at a constant emission rate (con.) and intermittent emissions at the threshold rate (int.).

| Basin | Strata | Observed | \% of sites with potential tank emissions > 3 g HC s^{-1} | | | | | | | | | |
| | | | water | | HC liquids (weighted HC EF) | | HC liquids (oil EF) | | HC liquids (condensate EF) | | total liquids (weighted HC EF) | |
			con.	int.	con.	int.	con.	int.	con.	int.	con.	int.
Bakken	Young	11.4%	0.0%	4.0%	52.5%	77.2%	51.8%	76.2%	89.6%	95.8%	52.5%	77.2%
	Old	12.8%	0.0%	3.3%	63.2%	87.2%	62.4%	86.9%	97.4%	98.4%	63.2%	87.2%
	Combined	12.2%	0.0%	3.6%	58.5%	82.8%	57.8%	82.2%	94.0%	97.3%	58.5%	82.8%
Barnett	High GOR	0.7%	0.1%	0.5%	0.1%	0.5%	0.0%	0.0%	0.0%	0.1%	0.1%	0.5%
	Medium GOR	1.4%	0.0%	0.9%	0.0%	1.5%	0.0%	0.5%	0.9%	3.9%	0.0%	1.5%
	Low GOR	19.3%	0.9%	4.3%	1.3%	10.1%	0.4%	4.9%	16.6%	31.7%	1.3%	10.1%
	Combined	3.3%	0.2%	1.1%	0.2%	2.0%	0.1%	0.8%	2.4%	5.3%	0.2%	2.0%
Eagle Ford	East	10.6%	0.0%	3.3%	30.3%	52.2%	29.9%	50.8%	83.0%	86.4%	30.3%	52.2%
	West	0.3%	0.0%	2.7%	23.0%	41.5%	20.9%	40.0%	73.5%	78.8%	23.0%	41.5%
	Combined	5.3%	0.0%	3.0%	26.5%	46.6%	25.2%	45.2%	78.0%	82.5%	26.5%	46.6%
Marcellus	Young High GOR	10.7%	0.0%	0.6%	9.7%	31.3%	2.9%	21.2%	50.5%	61.6%	9.7%	31.3%
	Old High GOR	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Low GOR	1.3%	0.0%	0.2%	0.0%	0.3%	0.0%	0.1%	0.2%	0.5%	0.0%	0.3%
	Combined	1.1%	0.0%	0.1%	0.5%	1.7%	0.1%	1.1%	2.6%	3.3%	0.5%	1.7%
Powder River	CBM	0.0%	0.0%	2.3%	0.0%	2.3%	0.0%	0.0%	0.6%	0.7%	0.0%	2.3%
	Oil/CBM	0.0%	0.0%	1.2%	0.3%	2.0%	0.3%	0.9%	8.0%	8.3%	0.3%	2.0%
	Oil	9.0%	0.0%	1.5%	6.0%	13.8%	6.0%	12.4%	64.9%	66.4%	6.0%	13.8%
	Combined	0.8%	0.0%	1.7%	0.6%	3.2%	0.6%	1.5%	9.5%	9.8%	0.6%	3.2%
Uintah	High GOR	1.4%	0.0%	0.2%	0.0%	1.6%	0.0%	1.2%	1.4%	5.5%	0.0%	1.6%
	Medium GOR	4.8%	0.0%	0.6%	0.0%	2.5%	0.0%	1.7%	0.4%	8.1%	0.0%	2.5%
	Low GOR	7.4%	0.0%	0.4%	0.0%	13.0%	0.0%	11.6%	15.5%	45.8%	0.0%	13.0%
	Combined	5.3%	0.0%	0.5%	0.0%	5.5%	0.0%	4.7%	5.0%	19.3%	0.0%	5.5%

Table S15. Sensitivity analysis of predicted percentage of well pads with controlled hydrocarbon (HC) emissions above 3 g HC s^{-1}. Emissions are calculated separately for produced water, HC liquids, and the combination of water and HCs using basin-level emission factors (EF) from the EPA O&G Estimation Tool 2014 version 1. Three different emission factors (EF) are used HC liquids: oil, condensate, and a weighted EF based on basin-level oil and condensate production. Controlled emissions are estimated by applying capture efficiency and control efficiency to a subset of highest emitting tanks based on basin-level control data from the EPA O&G Estimation Tool. For each emission estimate, the % of sites with emissions above the threshold is calculated based on two temporal profiles: continuous emissions at a constant emission rate (con.) and intermittent emissions at the threshold rate (int.).

Basin	Strata	Observed	HC liquids (weighted HC EF)		HC liquids (oil EF)		HC liquids (condensate EF)		total liquids (weighted HC EF)	
			con.	int.	con.	int.	con.	int.	con.	int.
Bakken	Young	11.4%	9.0%	31.1%	9.0%	31.1%	38.1%	67.0%	9.0%	31.8%
	Old	12.8%	6.5%	28.8%	6.5%	28.8%	35.2%	64.1%	6.8%	29.1%
	Combined	12.2%	7.6%	29.8%	7.6%	29.8%	36.5%	65.4%	7.8%	30.3%
Barnett	High GOR	0.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	Medium GOR	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.1%
	Low GOR	19.3%	0.0%	0.1%	0.0%	0.1%	0.0%	1.2%	0.0%	0.3%
	Combined	3.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.1%
Eagle Ford	East	10.6%	0.0%	2.9%	0.0%	2.9%	7.6%	27.2%	0.0%	3.0%
	West	0.3%	0.0%	1.9%	0.0%	1.9%	3.8%	20.6%	0.0%	2.1%
	Combined	5.3%	0.0%	2.4%	0.0%	2.3%	5.6%	23.8%	0.0%	2.5%
Marcellus	Young High GOR	10.7%	0.0%	1.0%	0.0%	0.6%	0.0%	6.8%	0.0%	1.0%
	Old High GOR	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Low GOR	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Combined	1.1%	0.0%	0.1%	0.0%	0.0%	0.0%	0.3%	0.0%	0.1%
Powder River	CBM	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.1%
	Oil/CBM	0.0%	0.0%	0.0%	0.0%	0.0%	1.3%	3.1%	0.0%	0.1%
	Oil	9.0%	0.0%	1.0%	0.0%	0.9%	22.4%	34.0%	0.0%	1.0%
	Combined	0.8%	0.0%	0.1%	0.0%	0.1%	2.5%	4.4%	0.0%	0.2%
Uintah	High GOR	1.4%	0.0%	0.3%	0.0%	0.3%	0.0%	1.3%	0.0%	0.5%
	Medium GOR	4.8%	0.0%	0.4%	0.0%	0.3%	0.0%	1.6%	0.0%	0.5%
	Low GOR	7.4%	0.0%	0.6%	0.0%	0.6%	0.0%	2.7%	0.0%	0.6%
	Combined	5.3%	0.0%	0.4%	0.0%	0.4%	0.0%	1.9%	0.0%	0.5%

Column group header: % of sites with controlled tank emissions > 3 g HC s^{-1}

Table S16. Summary of single parameter binomial generalized linear models predicting detection of hydrocarbon emissions based on well pad or operator regional parameters: GLM intercept and coefficient, Akaike Information Criterion, Pearson correlation (r) and significance (p) of fit between observed and predicted detection, Hosmer-Lemeshow goodness of fit (p > 0.05 indicates observed and predicted detection and not statistically different), and the ratio of predicted detection at the 97.5[th] and 2.5[th] percentile of parameter values.

Dataset	GLM Parameter	GLM intercept	GLM coefficient	AIC	GLM fit (r)	GLM fit (p)	Hosmer-Lemeshow (p)	97.5th / 2.5th ratio
All Basins	Pad Well Count	-3.741	3.39E-01	2651	0.123	3.3E-24	1.0E+00	3.2E+00
All Basins	Pad Well Age	-2.012	-1.65E-02	2530	0.205	1.4E-50	0.0E+00	8.9E-04
All Basins	Pad Gas Production	-3.250	1.06E-04	2709	0.092	2.1E-11	0.0E+00	1.3E+00
All Basins	Pad Oil Production	-3.323	1.39E-03	2629	0.174	5.4E-29	0.0E+00	1.9E+00
All Basins	Pad Water Production	-2.977	2.30E-04	2305	0.018	5.5E-03	1.1E-13	1.1E+00
All Basins	Pad % Energy from Oil	-3.910	2.08E+00	2520	0.168	1.1E-52	2.0E-04	6.0E+00
All Basins	Operator Well Count	-2.553	-3.76E-04	2646	0.108	2.6E-25	3.6E-13	1.2E-01
All Basins	Operator Gas Production	-2.978	-6.29E-07	2730	0.054	1.3E-06	0.0E+00	4.3E-01
All Basins	Operator Oil Production	-3.360	8.12E-06	2706	0.067	6.3E-12	0.0E+00	2.9E+00
All Basins	Operator Water Production	-2.648	-3.51E-06	2287	0.087	3.2E-07	2.8E-13	2.7E-01
All Basins	Operator % Energy From Oil	-3.944	2.20E+00	2560	0.164	6.6E-44	0.0E+00	5.4E+00
Fayetteville	Pad Well Count	-3.587	1.73E-01	109	0.074	2.2E-01	1.0E+00	2.8E+00
Fayetteville	Pad Well Age	-2.887	-3.98E-03	110	0.020	7.3E-01	3.1E-01	7.3E-01
Fayetteville	Pad Gas Production	-3.080	2.43E-06	111	0.001	9.9E-01	2.8E-01	1.0E+00
Fayetteville	Operator Well Count	-4.223	4.27E-04	108	0.082	1.3E-01	NaN	2.7E+00
Fayetteville	Operator Gas Production	-4.107	6.25E-07	108	0.083	1.3E-01	NaN	2.7E+00
Barnett	Pad Well Count	-3.871	3.14E-01	504	0.055	4.9E-04	1.0E+00	2.4E+00
Barnett	Pad Well Age	-1.975	-1.34E-02	471	0.258	1.8E-11	0.0E+00	2.3E-04
Barnett	Pad Gas Production	-3.616	6.04E-04	498	0.117	2.6E-05	1.0E-01	3.0E+00
Barnett	Pad Oil	-3.903	7.30E-02	387	0.445	7.1E-30	7.7E-04	5.9E+00

106

	Production						
Barnett	Pad Water Production	-3.340	1.47E-04	514	0.014	1.5E-01	1.4E-12	1.1E+00
Barnett	Pad % Energy from Oil	-3.669	2.71E+00	470	0.173	1.2E-11	9.1E-01	7.1E+00
Barnett	Operator Well Count	-2.327	-3.83E-04	472	0.132	3.7E-11	2.9E-05	1.2E-01
Barnett	Operator Gas Production	-2.328	-1.57E-06	471	0.132	2.1E-11	NaN	1.2E-01
Barnett	Operator Oil Production	-4.429	2.41E-04	446	0.288	6.2E-17	7.2E-11	1.1E+01
Barnett	Operator Water Production	-4.583	6.54E-06	487	0.238	5.9E-08	0.0E+00	9.1E+00
Barnett	Operator % Energy From Oil	-3.608	2.64E+00	494	0.114	2.2E-06	NaN	2.6E+00
Eagle Ford	Pad Well Count	-3.300	1.81E-01	231	0.099	1.7E-02	1.0E+00	2.7E+00
Eagle Ford	Pad Well Age	-2.247	-2.15E-02	228	0.125	2.8E-03	1.3E-01	9.0E-06
Eagle Ford	Pad Gas Production	-3.376	4.07E-04	209	0.313	1.4E-07	6.5E-01	4.5E+00
Eagle Ford	Pad Oil Production	-3.402	9.93E-04	205	0.361	1.7E-08	4.7E-01	4.9E+00
Eagle Ford	Pad Water Production	-3.061	1.07E-03	230	0.118	8.4E-03	6.9E-01	2.7E+00
Eagle Ford	Pad % Energy from Oil	-2.819	-5.56E-02	237	0.003	9.3E-01	1.9E-02	9.5E-01
Eagle Ford	Operator Well Count	-2.764	-1.59E-04	237	0.015	7.1E-01	0.0E+00	7.8E-01
Eagle Ford	Operator Gas Production	-3.418	2.19E-06	234	0.068	8.2E-02	2.8E-07	2.4E+00
Eagle Ford	Operator Oil Production	-3.148	3.18E-06	236	0.044	2.4E-01	0.0E+00	2.0E+00
Eagle Ford	Operator Water Production	-2.823	-7.30E-07	237	0.013	7.7E-01	0.0E+00	8.2E-01
Eagle Ford	Operator % Energy From Oil	-2.599	-4.02E-01	237	0.015	6.8E-01	1.3E-09	7.5E-01
Bakken	Pad Well Count	-2.630	4.86E-01	523	0.223	1.2E-07	1.0E+00	2.4E+00
Bakken	Pad Well Age	-1.577	-5.59E-03	544	0.111	1.1E-02	2.7E-01	1.4E-01
Bakken	Pad Gas Production	-2.109	7.18E-04	526	0.211	6.1E-07	8.8E-02	2.6E+00
Bakken	Pad Oil Production	-2.122	8.57E-04	528	0.195	1.9E-06	2.7E-01	2.6E+00

Bakken	Pad Water Production	-1.927	5.50E-04	546	0.079	3.2E-02	4.5E-02	1.6E+00
Bakken	Pad % Energy from Oil	-1.015	-9.96E-01	550	0.031	3.0E-01	4.1E-01	7.4E-01
Bakken	Operator Well Count	-1.183	-1.08E-03	534	0.155	4.3E-05	4.3E-01	2.8E-01
Bakken	Operator Gas Production	-1.267	-6.74E-06	538	0.132	2.5E-04	3.7E-01	3.4E-01
Bakken	Operator Oil Production	-1.150	-9.65E-06	535	0.144	6.0E-05	9.9E-01	2.8E-01
Bakken	Operator Water Production	-1.283	-1.02E-05	534	0.150	3.3E-05	7.8E-01	2.7E-01
Bakken	Operator % Energy From Oil	-1.662	-2.06E-01	551	0.003	9.5E-01	4.9E-03	9.8E-01
Marcellus	Pad Well Count	-6.080	6.73E-01	182	0.376	7.3E-20	1.0E+00	1.2E+01
Marcellus	Pad Well Age	-0.901	-5.76E-02	176	0.307	3.0E-21	3.7E-03	1.6E-08
Marcellus	Pad Gas Production	-4.774	1.20E-04	226	0.270	2.8E-10	8.8E-05	1.9E+00
Marcellus	Pad Oil Production	-4.509	7.11E-03	261	0.016	3.9E-02	1.0E+00	1.3E+00
Marcellus	Pad % Energy from Oil	-4.559	2.66E+00	259	0.055	8.4E-03	7.0E-01	1.7E+00
Marcellus	Operator Well Count	-6.166	4.55E-04	254	0.089	7.8E-04	6.9E-04	1.2E+01
Marcellus	Operator Gas Production	-5.922	2.76E-06	229	0.109	1.4E-09	2.2E-01	1.1E+01
Marcellus	Operator Oil Production	-4.971	1.39E-04	255	0.077	1.2E-03	1.2E-03	3.9E+00
Marcellus	Operator % Energy From Oil	-4.689	5.73E+00	262	0.038	6.2E-02	3.3E-01	3.0E+00
Uintah	Pad Well Count	-2.926	2.44E-01	680	0.024	2.1E-01	1.0E+00	1.2E+00
Uintah	Pad Well Age	-1.858	-9.06E-03	649	0.152	1.5E-08	1.0E-03	1.6E-02
Uintah	Pad Gas Production	-2.723	1.01E-03	678	0.042	5.8E-02	7.3E-01	1.4E+00
Uintah	Pad Oil Production	-2.878	2.41E-02	664	0.124	3.8E-05	4.8E-05	2.6E+00
Uintah	Pad Water Production	-2.677	1.39E-03	681	0.016	4.4E-01	2.7E-03	1.3E+00
Uintah	Pad % Energy from Oil	-3.006	8.22E-01	671	0.083	1.2E-03	1.5E-01	2.1E+00
Uintah	Operator	-2.159	-4.33E-	673	0.064	4.5E-03	NaN	3.2E-01

108

	Well Count		04					
Uintah	Operator Gas Production	-2.306	-3.74E-06	664	0.092	3.1E-05	8.5E-01	1.2E-01
Uintah	Operator Oil Production	-2.741	1.09E-05	680	0.029	2.9E-01	NaN	1.3E+00
Uintah	Operator Water Production	-2.059	-2.25E-05	669	0.082	3.7E-04	2.7E-01	2.2E-01
Uintah	Operator % Energy From Oil	-2.903	6.41E-01	677	0.052	4.9E-02	NaN	1.6E+00
Powder River	Pad Well Count	-5.817	1.06E+00	169	0.034	3.9E-02	1.0E+00	2.8E+00
Powder River	Pad Well Age	-1.624	-7.76E-02	127	0.263	1.3E-11	0.0E+00	3.8E-12
Powder River	Pad Gas Production	-5.056	1.99E-03	168	0.053	2.2E-02	3.5E-01	3.4E+00
Powder River	Pad Oil Production	-4.779	3.02E-03	158	0.069	1.5E-04	2.0E-08	1.1E+00
Powder River	Pad Water Production	-4.701	9.33E-04	173	0.009	5.8E-01	7.6E-01	1.5E+00
Powder River	Pad % Energy from Oil	-6.340	4.37E+00	123	0.193	1.4E-12	9.4E-01	4.5E+01
Powder River	Operator Well Count	-2.700	-2.04E-03	144	0.126	6.1E-08	NaN	1.2E-02
Powder River	Operator Gas Production	-2.985	-2.01E-05	142	0.129	2.3E-08	7.5E-02	3.6E-03
Powder River	Operator Oil Production	-6.505	3.18E-04	150	0.128	1.3E-06	8.5E-01	2.5E+01
Powder River	Operator Water Production	-2.951	-2.96E-05	140	0.135	1.1E-08	2.9E-02	2.1E-03
Powder River	Operator % Energy From Oil	-7.839	7.46E+00	124	0.184	3.4E-12	NaN	1.0E+02

Table S17. Binomial generalized linear model A4 coefficients, standard errors, and p values for model terms and interactions. Model terms are summarized in Table S5

	Coefficients	Standard Error	p value
basin (Bakken)	-1.61E+00	9.45E-01	0.089
basin (Barnett)	-3.56E+00	4.05E-01	<2E-16
basin (Eagle Ford)	-2.18E+00	7.33E-01	0.003
basin (Fayetteville)	-3.54E+00	1.27E+00	0.005
basin (Marcellus)	-2.94E+00	7.91E-01	0.000
basin (Powder River)	-4.79E+00	1.46E+00	0.001
basin (Uintah)	-2.35E+00	3.34E-01	0.000
well count	2.99E-01	1.26E-01	0.018
well age	-1.67E-03	2.05E-03	0.417
gas production	2.48E-04	2.97E-04	0.404
oil production	1.44E-04	3.88E-04	0.711
percent energy from oil	-9.56E-01	1.08E+00	0.377
Barnett: well count	-2.59E-01	2.20E-01	0.240
Eagle Ford: well count	-6.76E-01	2.44E-01	0.006
Fayetteville: well count	5.12E-02	2.41E-01	0.832
Marcellus: well count	1.78E-01	1.70E-01	0.293
Powder River: well count	-1.85E+00	9.17E-01	0.044
Uintah: well count	-5.12E-01	2.47E-01	0.038
Barnett: well age	-2.35E-03	2.81E-03	0.403
Eagle Ford: well age	-7.79E-03	8.75E-03	0.373
Fayetteville: well age	-1.46E-04	1.75E-02	0.993
Marcellus: well age	-4.07E-02	1.21E-02	0.001
Powder River: well age	-2.42E-02	9.80E-03	0.014
Uintah: well age	-4.32E-03	2.74E-03	0.116
Barnett: gas production	-2.35E-04	4.15E-04	0.571
Eagle Ford: gas production	-8.23E-05	3.36E-04	0.807
Fayetteville: gas production	-5.84E-04	4.45E-04	0.190
Marcellus: gas production	-2.80E-04	2.99E-04	0.348
Powder River: gas production	2.06E-03	9.66E-04	0.033
Uintah: gas production	9.64E-04	6.71E-04	0.151
Barnett: oil production	5.11E-02	9.98E-03	0.000
Eagle Ford: oil production	1.11E-03	6.36E-04	0.080
Fayetteville: oil production	NA	NA	NA
Marcellus: oil production	-1.08E-02	5.30E-03	0.042
Powder River: oil production	3.66E-06	1.48E-03	0.998
Uintah: oil production	9.19E-03	6.70E-03	0.170
Barnett: percent energy from oil	2.67E+00	1.24E+00	0.032
Eagle Ford: percent energy from oil	2.86E-01	1.46E+00	0.844
Fayetteville: percent energy from oil	NA	NA	NA
Marcellus: percent energy from oil	3.81E+00	1.58E+00	0.016
Powder River: percent energy from oil	8.08E+00	1.70E+00	0.000
Uintah: percent energy from oil	1.66E+00	1.14E+00	0.145

Figure S1. Areas including surveyed grid cells in seven basins: 1) Bakken, 2) Barnett, 3) Eagle Ford, 4) Fayetteville, 5) Marcellus, 6) Powder River, and 7) Uintah. Locations of individual surveyed pads are provided in the SI file "surveyed_well_pads.csv". Base imagery is from Google Earth (Map data: SIO, NOAA, U.S. Navy, NGA, GEBCO; © 2016 Google; Image Landsat, US Dept of State Geographer)
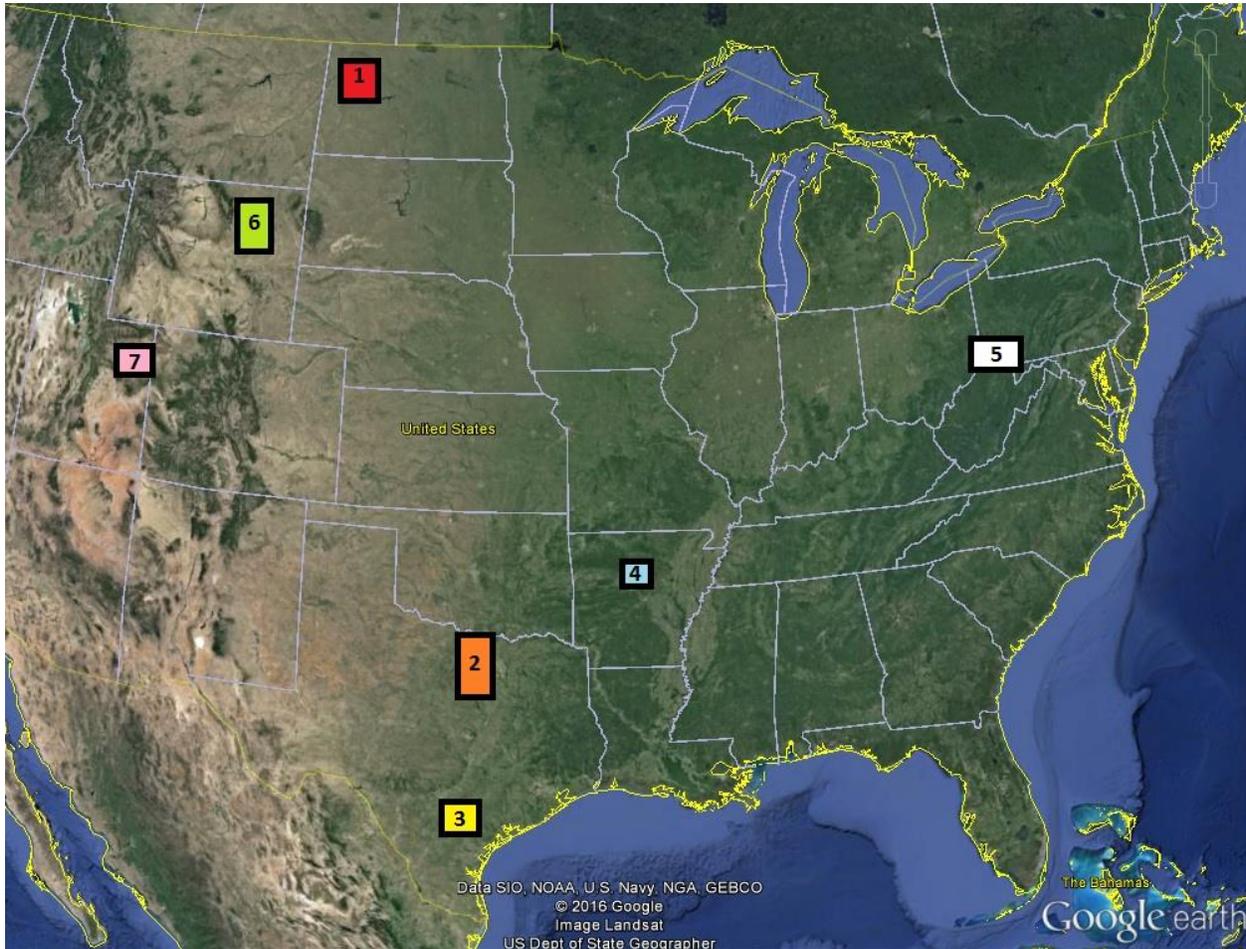
Figure S2. Percentage of well pads with detected emissions by deciles of operator parameters: a) Regional Well Count, b) Regional Gas Production (Mcf/day), c) Regional Oil Production (bbl/day), d) Regional Water Production (bbl/day), and e) Regional % Energy from Oil. The median values of each decile are displayed on the x-axes.
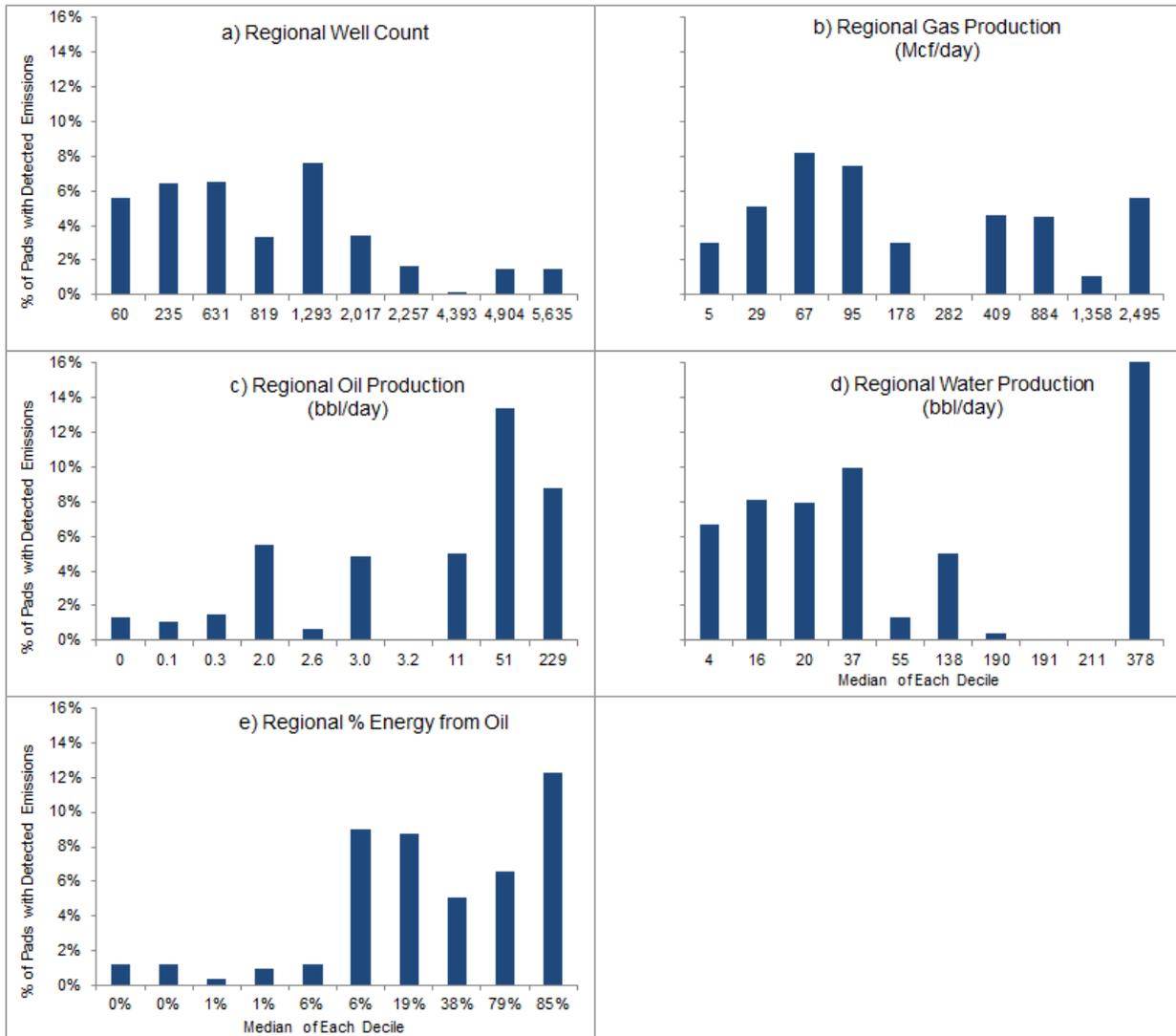
CHAPTER 3 APPENDIX REFERENCES

1.   Conley SA, Faloona IC, Lenschow DH, Karion A, Sweeney C. A Low-Cost System for Measuring Horizontal Winds from Single-Engine Aircraft. J Atmospheric Ocean Technol. 2014 Jun;31(6):1312–20.

2.   Conley SA, Faloona I, Miller GH, Lenschow DH, Blomquist B, Bandy A. Closing the dimethyl sulfide budget in the tropical marine boundary layer during the Pacific Atmospheric Sulfur Experiment. Atmospheric Chem Phys. 2009 Nov 17;9(22):8745–56.

3.   Safitri A, Gao X, Mannan MS. Dispersion modeling approach for quantification of methane emission rates from natural gas fugitive leaks detected by infrared imaging technique. J Loss Prev Process Ind. 2011 Mar;24(2):138–45.

4.   Lavoie TN, Shepson PB, Cambaliza MOL, Stirm BH, Conley S, Mehrotra S, Faloona IC, Mayfield M, Garcia K, Lyon D, et al. Temporal Variability of Methane Emissions at Oil and Natural Gas Operations in the Eagle Ford Basin. (In preparation for submission to Environmental Science & Technology)

5.   Greenhouse Gas Reporting Program [Internet]. Washington, DC: United States Environmental Protection Agency; 2015. Available from: http://www.epa.gov/ghgreporting

6.   Iowa Environmental Mesonet ASOS-AWOS-METAR Data Download [Internet]. Ames, IA. Available from: https://mesonet.agron.iastate.edu/request/download.phtml

7.   Oil and Gas Emission Estimation Tool 2014 Version 1 [Internet]. Washington, DC: United States Environmental Protection Agency; 2015. Available from: ftp://ftp.epa.gov/EmisInventory/2011nei/doc/Tool_and_Report112614.zip

CHAPTER 4

Site-level Gaussian dispersion model to optimize the deployment of continuous methane sensors

Introduction

Methane emissions from the oil and gas (O&G) supply chain are dominated by a relatively small fraction of facilities with high emission rates.[1–3] In Lyon et al. (2015) and Zavala-Araiza et al. (2016), accounting for super-emitters contributed to the reconciliation of top-down and bottom-up estimates of methane emissions in the Barnett Shale.[4,5] Although the term "super-emitter" typically refers to facility-level emissions, components also have highly skewed emission rate distributions. In both a national study of 150 natural gas well pads and a local study of 375 well pads in Fort Worth, TX, the top 20% of equipment leaks contributed over 80% of total emissions.[1,6] Vented sources such as pneumatic controllers and tanks that emit during normal operation may also have skewed distributions caused by malfunctioning devices. For example, Allen et al. (2015) report that 19% of 377 measured pneumatic controllers were responsible for 95% total emissions.[7] Facilities with high emissions may have a small number of components causing the vast majority of their emissions. The two transmission and storage facilities classified as super-emitters in Subramanian et al. (2015) had leaking compressor isolation valves with extremely high emission rates.[3] In a national study of 114 gathering stations, the 20% of sites with substantial tank venting had on average four times the emission rate of facilities without observed venting.[2] Based on a helicopter-based infrared camera survey of over 8,000 well pads, Lyon et al. (2016) found that over 90% of high emission sources were from storage tanks and some of these sites had ineffective tank control systems.[8] These high emission sources had weak relationships with site characteristics and the best statistical model could only 14% of the variance; therefore, frequent monitoring of all sites is required to quickly detect the occurrence of super-emitters.[8]

The rapid identification and mitigation of equipment leaks and malfunctioning equipment with high emission rates is critical to reducing emissions. Traditional leak detection methods such as Method 21 (measuring methane concentrations directly at individual components) and optical gas imaging (OGI) are rarely used more frequently than monthly at well pads due to the high labor and travel cost of sending

staff to widely distributed sites. The United States Environmental Protection Agency (US EPA) has

proposed a rule that would require new sites to be inspected with OGI semi-annually initially with a

change to quarterly or annually if the number of leaks is higher than 3% or lower than 1%, respectively.[9]

OGI performed on the ground by a skilled technician is highly effective and can typically detect leaks of

60 g hr^{-1} or larger.[10] Due to the skewed distribution of emission rates, a leak detection and repair (LDAR)

program that has a higher detection limit for identifying leaks but faster detection and repair time may lead

to greater emission reductions than OGI. Based on the leak distribution from Allen et al. (2014)[1], an

LDAR program with a 60 g hr^{-1} detection limit and semi-annual frequency to identify and repair leaks

would reduce annual emissions by 49%, while an alternative program with a 1000 g hr^{-1} detection limit but

weekly frequency would reduce emissions by 53%.

The key to cost-effectively enhancing the ability of LDAR programs to reduce emissions is to increase

inspection frequency while minimizing labor costs by targeting staff deployment to high emission sites.

One option is to develop LDAR programs that utilize mobile or aerial surveys to inspect multiple sites

without requiring staff to visit individual sites. A recent paper proposed a method for using work trucks

with methane sensors coupled with inversion dispersion modeling to efficiently survey large numbers of

sites.[11] Several companies offer OGI from helicopters of fixed-wing aircraft to detect large leaks.[12,13]

Other companies are developing leak detection systems deployed on unmanned aerial vehicles.[14,15]

Satellite remote sensing has also been used to quantify O&G methane emissions, but the resolution of

existing satellites is insufficient for site-level measurements.[16]   Another option for frequent LDAR is to

deploy low-cost, continuous sensors at O&G sites. A previous study has used dispersion modeling at the

field-level to demonstrate the cost-effectiveness of several alternative LDAR programs including

unmanned aerial vehicles and continuous sensors. [17] The study estimated that LDAR programs with a

positive net present value from recovered gas could mitigate 80% of leaks and their cost-effectiveness

could be improved by targeting the largest leaks.[17]

Existing commercially available methane analyzers are either too expensive or not sensitive enough to be

used for continuous monitoring of O&G sites. There are two ongoing technology development programs

that are working to accelerate the availability of low cost monitors. The United States Department of

Energy Advanced Research Projects Agency – Energy (ARPA-E) MONITOR program has funded eleven companies developing low-cost technologies to detect O&G methane leaks.[18] The goal of the program is to "detect and measure methane leaks as small as 1 ton per year from a site 10 m x 10 m area with a certainty that would allow 90% reduction in methane loss for an annual site cost of $3,000."[18] This includes detecting emissions ≥ 6 standard cubic feet per hour (114 g hr^{-1}), locating the leak to within 1 m, and quantifying the leak within 20% error. The ARPA-E technologies are primarily in early stages of development and likely will not be ready for commercialization until after the program concludes in 2018. Environmental Defense Fund has sponsored another program, the Methane Detectors Challenge (MDC), which is designed to catalyze the development of first generation low-cost continuous leak detection systems.[19] The MDC goal is to detect a 150 scfh leak (2,880 g hr^{-1}) for less than $1,000 per site. The system is not required to quantify or locate the leak, but should be able to distinguish onsite leaks from either offsite sources or onsite vented sources such as pneumatic controllers. Systems must be able to operate with minimal maintenance under diverse environmental conditions including temperatures from -20 to 135^0 F. Two MDC technologies performed sufficiently well during two phases of laboratory and outdoor testing to advance to pilot deployments at partner company sites, which are expected to commence in summer 2016. Both technologies utilize tunable-diode lasers that are methane-specific: the first, manufactured by Quanta3, is an enclosed system that measures methane concentration at a point, while the second, manufactured by Dalian Actech, employs an open path configuration to measure the path-averaged concentration between the laser and a retro-reflector.

For this work, I have developed a leak dispersion model that simulates O&G sites to optimize the deployment of point and open path monitors by estimating the time between detection based on site layout, emission rates, sensor detection limit, and local meteorological data. Model results will assist operators participating in the MDC to deploy point and open path systems in positions that will most quickly detect leaks. The model also demonstrates the relationship of sensor detection limits in concentration enhancement to source emission rates, which can guide improvements of leak detection systems for identifying smaller leaks. This chapter describes the model, sensitivity of detection to source-receptor orientation and distance based on three sets of meteorological data, and results of model runs based on three O&G well pads.

Methods

Methane emissions are simulated with a custom Gaussian dispersion model programmed by the author in the open source language R.[20] Gaussian dispersion assumes a constant emission rate, an infinite plume, negligible downwind dispersion, no wind shear, non-reactive pollutants, and constant meteorological conditions during the modeled period.[21] Although O&G emission may violate some assumptions of Gaussian dispersion, the model is valuable as a first approximation of concentration enhancements at multiple receptors from potential component-level emission sources.

O&G sites are defined by a Cartesian coordinate system with the origin at a site corner. The x-axis and y-axis either are oriented east and north, respectively, or the degrees offset from this orientation are reported. All potential emission sources are included in a comma-separated value file (sources.csv) including source identification code (*source.id*), x, y, and z coordinates in meters from the origin (*source.x*, *source.y*, and *source.z*), and source type category such as equipment leak or pneumatic controller (*source.type*). Potential locations for point sensors are included in a separate file (receptors.csv) including receptor identification code (*receptor.id*), x, y, and z coordinates in meters from the origin (*receptor.x*, *receptor.y*, and *receptor.z*). Potential locations for open path sensors are included in the same file by defining a group of point receptors with a separate path identification code (*path.id*); in the current build of the model, the receptors must be evenly spaced for accurate estimation of path-averaged concentration. Meteorological data are includes as a file (met.csv) with ten years of hourly wind speed in m s^{-1} (*u*), wind direction in degrees (*v*), stability class as A, B, C, D, E, or F (*stab*); the data are identified with a sequential number from the start of the time series (*hour.id*). An additional file (Caraway.csv) includes power law variables by stability class used to estimate dispersion parameters. Prior to running the model, the user sets methane emission rates for each source type in standard cubic feet per hour (scfh), three leak detection thresholds in parts per million (ppm) above background, and the minimum wind speed for which concentrations are calculated.

The model uses ten years of hourly meteorological data based on airport weather observations (METAR). In a separate spreadsheet model, Pasquill stability classes are estimated using a modified version of Turner (1969) based on wind speed, solar radiation, and cloudiness.[22,23,21] Solar zenith angles are

117

calculated based on time and latitude/longitude using a publically available spreadsheet calculator.[24]
Solar radiation values are grouped into three categories (slight, moderate, strong) based on solar angle class ($0–15^0$, $15–35^0$, $35–60^0$, $>60^0$) adjusted by cloud cover class (≤50%, >50% and ≥7000 ft ceiling, >50% and <7000 ft ceiling). For daytime hours, stability class is based on wind speed class (<2, 2–3, 3–5, 5–6, >6 m s^{-1}) and solar radiation class. For nighttime hours, stability class is based on wind speed class and cloud cover class. Intermediate stability classes are replaced with the most stable class (e.g., A for A-B); a stability class of F was used for low wind speed, nighttime conditions. Missing data are replaced with the closest available values from preceding or following hours.

The model's initial step is to calculate for each source the hourly concentration enhancement of methane in ppm above background at every point receptor for every hour of meteorological data. First, downwind and crosswind distance of each receptor from a source is calculated based on their Cartesian coordinates and wind direction using equations 1 and 2. If site axes are not oriented north/east, then hourly wind direction is adjusted by the degrees offset. Second, horizontal and vertical plume dispersion coefficients (σy, σz) are calculated using equations 3 and 4 based on downwind distance and Caraway's power law exponents for the stability class.[21] Methane concentration (g m^{-3}) is calculated using the Gaussian dispersion equation (equation 5) based on wind speed, σy, and σz and a source type-specific emission rate, then converted to ppm assuming conditions of 1 atmosphere and 25^0 C (1 g m^{-3} = 1,527 ppm). Emission rates are reported in standard cubic feet per hour (scfh), the most commonly used unit by US operators (1 scfh = 19.2 g hr^{-1}). If hourly wind speed is below the set threshold, then the concentration is set as not available to avoid the high uncertainty of the Gaussian dispersion equation and simulate the likely restriction of sensor alarms during low wind speed conditions when emissions can pool near sources.

Equations 1 & 2

$$dw = \left(-(receptor.x - source.x) * \sin\left(v * \frac{\pi}{180}\right)\right) - ((receptor.y - source.y * \cos(v * \pi/180))$$

$$cw = \left((receptor.x - source.x) * \cos\left(v * \frac{\pi}{180}\right)\right) - ((receptor.y - source.y * \sin(v * \pi/180))$$

where,

dw = downwind distance of receptor from source (m)

cw = crosswind distance of receptor from source (m)

receptor.x, receptor.y = Cartesian coordinates of receptor in meters

source.x, source.y = Cartesian coordinates of source in meters

v = wind direction (^{0})

Equation 3 & 4

$$\sigma_z = a * dw^b$$

$$\sigma_y = c * dw^d$$

where,

σ_z = vertical plume dispersion coefficient (m)

σ_y = horizontal plume dispersion coefficient (m)

dw = downwind distance of receptor from source (m)

a, b, c, d = constants specific to each stability class (A, B, C, D, E, F)

Equation 5

$$C = \frac{Q}{2\pi * u * \sigma_y * \sigma_z} * e^{-\frac{cw^2}{2\sigma_y^2}} * \left(e^{-\frac{(receptor.z - source.z)^2}{2\sigma_z^2}} + e^{-\frac{(receptor.z + source.z)^2}{2\sigma_z^2}} \right)$$

where,

C = concentration (g m^{-3})

u = wind speed (m s^{-1})

σz = vertical plume dispersion coefficient (m)

σy = horizontal plume dispersion coefficient (m)

cw = crosswind distance of receptor from source (m)

receptor.z = receptor height relative to Cartesian coordinate system origin (m)

source.z = source height relative to Cartesian coordinate system origin (m)

After calculating hourly concentration enhancement at each receptor caused by a single source's emissions, the model flags which hours are above three set detection limits. Next, the model calculates the fraction of hours in which emissions at the set rate and detection limits would be detected, the average number of hours between detection during the ten years of meteorological data, and the 95th percentile hours between detection. For open path receptors, the model first calculates the average hourly concentration by path to estimate the path-averaged concentrations, and then applies the same steps to determine the hours between detection. Once the model calculates the results for each source, the best receptor location is selected based on the lowest median value of individual sources' average hours between detection. This represents the average time expected to identify an emission source after it starts based on the site layout and local meteorological data. Site-level performance can be based on either all sources or the subset of equipment leak sources depending on if the system is also intended to detect emissions from vented sources. For the best receptor locations, average hours between detection are compared for each source to determine which sources have the shortest and longest detection time.

Model Runs

The model was run for three example sites in different oil and gas basins: the Eagle Ford Shale (south Texas), Fayetteville Shale (north-central Arkansas, and Bakken Shale (west North Dakota). Meteorological data were based on January 1, 2006 to December 31, 2015 hourly data at the Lackland Airforce Base (SFK), Little Rock Airforce Base (LRF), and Sloulin Field International Airport (ISN) for the Eagle Ford, Fayetteville, and Bakken, respectively.[25] The minimum wind speed was set at 0.2 m s^{-1}. For

all scenarios, the primary leak detection system sensitivity was set at 0.5 ppm above background and alternative limits were set at 0.1 ppm and 1 ppm. For open path sensors, the detection limit is the path-averaged concentration enhancement. These values are based on developers' initial estimates of detection limits for the two MDC leak detection systems.

Sensitivity

For all three sets of meteorological data, the model was run using a source and receptor setup designed to test the sensitivity of detection to orientation, distance, and relative height of the receptor from the sources. The site has a single source located at the origin and a series of receptors oriented along sixteen 22.5^0 spaced radial lines located from 1 to 1000 m from the source. At 10 and 100 m distance, receptors are located from -5 to 5 m height at 1 meter intervals.

Eagle Ford Shale

Statoil gas well pad Neller Gas Unit 1 located in Karnes County, TX (28.874^0, -97.730^0) was selected for the first example site. This site is offset from a north orientation by 67^0; the site origin is defined as the fence line northwest corner with the x- and y-axes oriented 157^0 and 67^0, respectively. The site fence line is 138 m along the x-axis and 158 m along the y-axis. Based on Google Earth imagery, the site includes 2 wellheads, 2 separators, 1 meter, three pneumatic controllers, and approximately 72 connectors (locations estimated at 3 m intervals along aboveground pipelines). Cartesian coordinates were estimated based on Google Earth imagery (Figures 1 & 2). Sources were assumed to be at a height of 1 m with the exception of separators, which had three leak points at 1, 2, 3 m heights. Receptors were modeled at 2 m intervals along the fence line plus an interior line parallel to the y-axis at a 10 m distance from the closest source. All receptors were at a height of 1 m. Paths were defined as the north, east, south, and west fence lines (name corresponding the offset adjusted orientation) and interior path. For the first model run, emissions were modeled for all sources at 150 scfh. For the second model run, emissions were modeled for leak sources at 6 scfh.

Figure 3. Eagle Ford shale well pad used for model scenarios. Imagery is from Google Earth.

Figure 4. Eagle Ford shale well pad source (black) and receptor (red) locations used for model scenarios. Pneumatic controllers are highlighted in green. Path names are shown in purple.
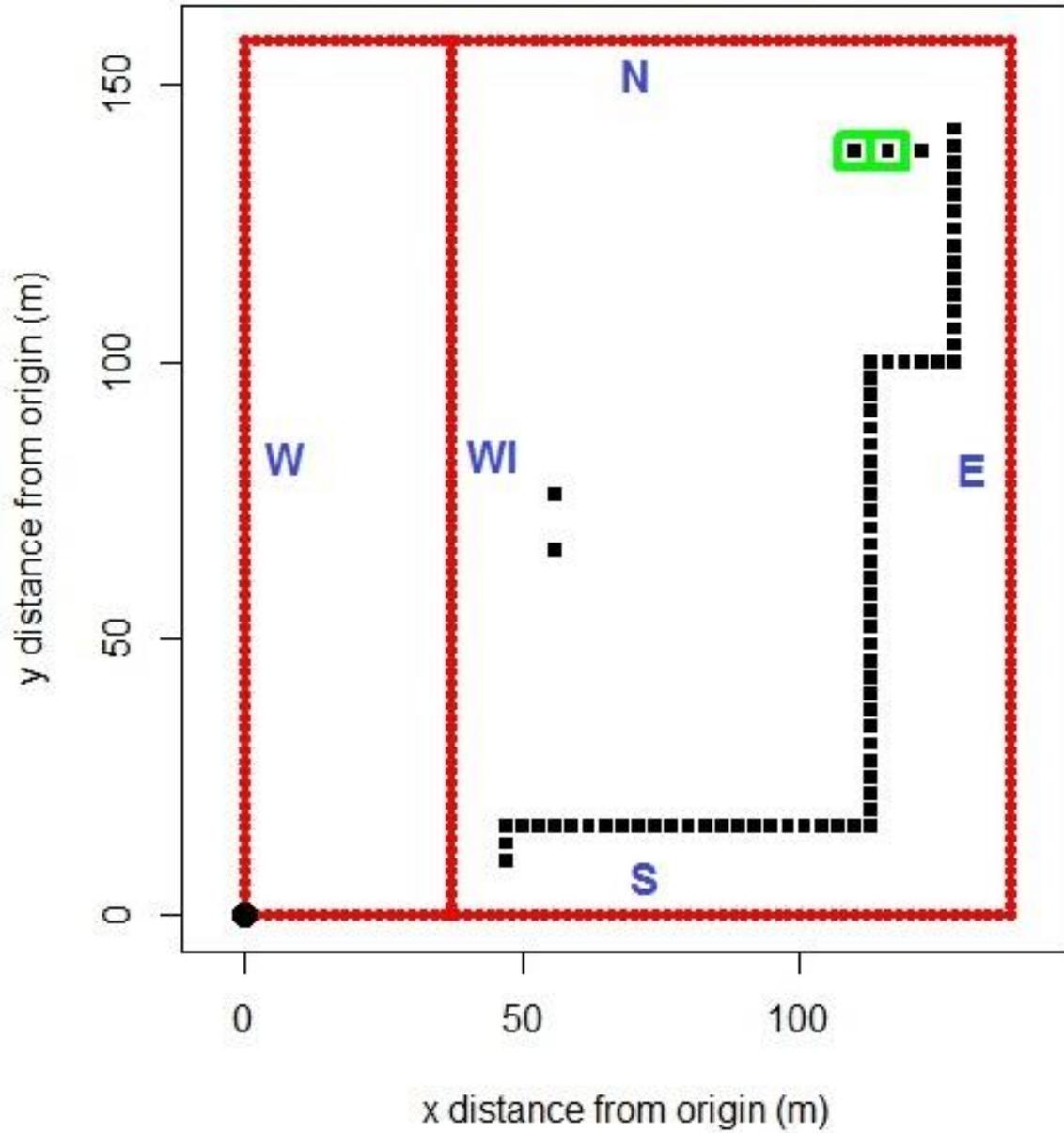
Fayetteville

Southwestern Energy gas well pad Hutchins 09-13 located in Van Buren County, AR (35.409^0, -92.369^0) was selected for the second example site. The site origin is defined as the fence line southwest corner with the x- and y-axes oriented east and north, respectively. The site fence line is 200 m along the x-axis and 180 m along the y-axis. The site includes five wellheads, five chemical injection pumps, five separators, five meters, three produced water tanks, and approximately 41 connectors (locations estimated at 3 m intervals along aboveground pipelines). Cartesian coordinates were estimated based on Google Earth imagery and operator-provided information (Figures 3 & 4). Most sources were assumed to be at a height of 1 – 1.5 m. For separators, three leak points were modeled at 0.5, 1.5, and 2.5 m. For tanks, emissions were modeled from hatches at 6 m height. Receptors were modeled at 2 m intervals along the fence line plus at 10 m intervals along two interior lines parallel to the x-axis at a distance of 6 and 16 m from the closest source. All receptors were at a height of 1 m. Paths were defined as the north, east, south, and west fence lines and two interior paths. For the first model run, emissions were modeled for all sources at 150 scfh. For the second model run, emissions were modeled for leak sources at 6 scfh.

Figure 5. Fayetteville shale well pad used for model scenarios. Imagery is from Google Earth.

Figure 6. Fayetteville shale well pad source (black) and receptor (red) locations used for model scenarios. Vented sources are highlighted in green, purple, and blue for pneumatic controllers, chemical injection pumps, and produced water tanks, respectively. Path names are shown in purple.
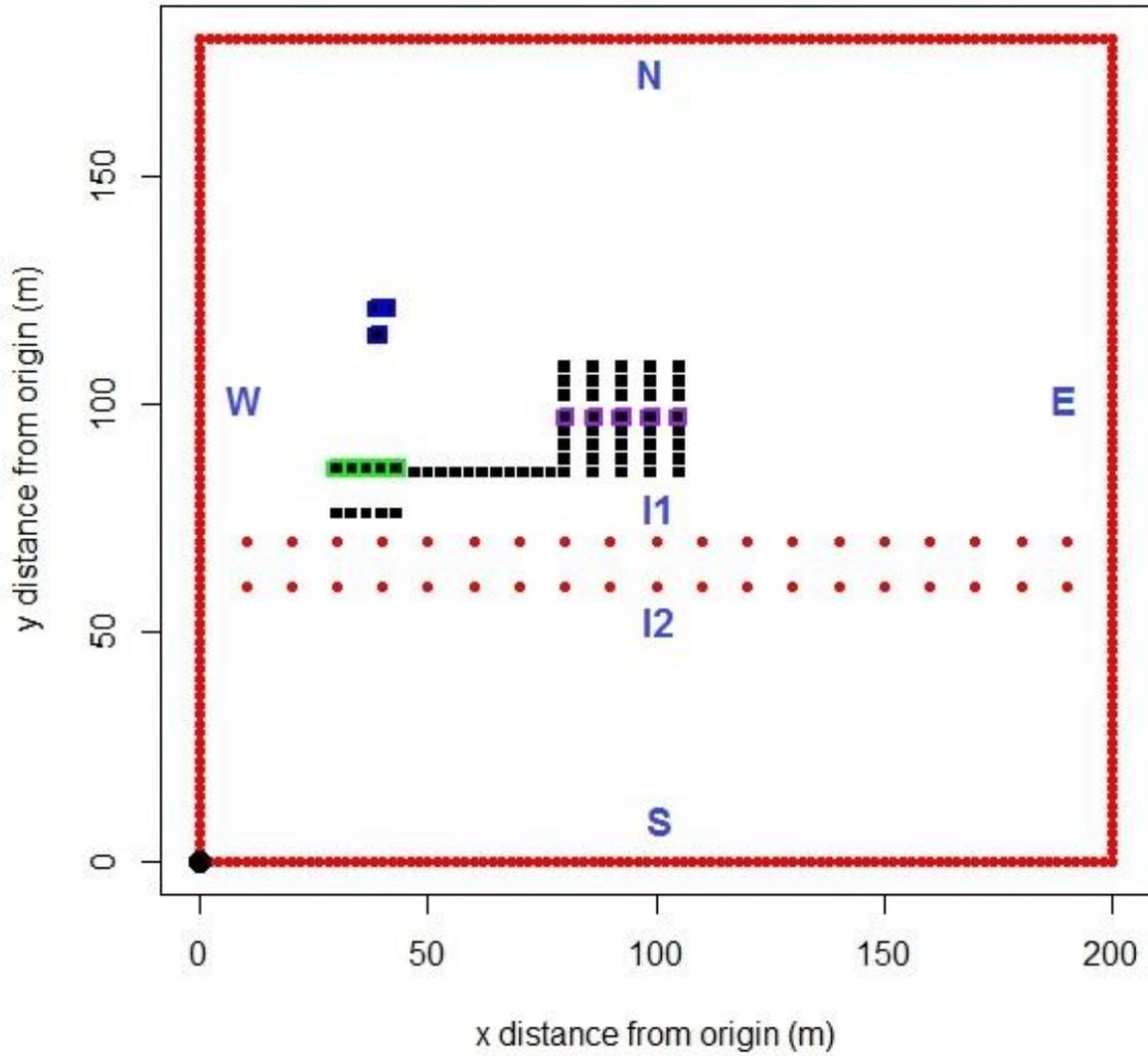
Bakken

Statoil oil well pad Lippert 1-12 1-H located in Williams County, ND (48.110, -103.746) was selected for the third example site. The site origin is defined as the fence line southwest corner with the x- and y-axes oriented east and north, respectively. The site fence line is 116 m along the x-axis and 136 m along the y-axis. Based on Google Earth imagery, the site includes one pump jack, one separator, one pneumatic controller, one heater, one produced water tank, twelve oil tanks, and approximately 18 connectors (locations estimated at 3 m intervals along aboveground pipelines). Cartesian coordinates were estimated based on Google Earth imagery (Figures 5 & 6). Most sources were assumed to be at a height of 1 m. For separators, three leak points were modeled at 2, 4, and 6 m. For tanks, emissions were modeled from hatches at 6 m height. Receptors were modeled at 2 m intervals along the fence line at heights of both 1 m and 5 m. Paths were defined as the north, east, south, and west fence lines at the two different heights. For the first model run, emissions were modeled for all sources at 150 scfh. For the second model run, emissions were modeled for leak sources at 6 scfh and vented emissions at 4.9, 4.4, and 83 scfh for pneumatic controllers, produced water tanks, and oil tanks, respectively.

Figure 7. Bakken shale well pad used for model scenarios. Imagery is from Google Earth.

Figure 8. Bakken shale well pad source (black) and receptor (red) locations used for model scenarios. Vented sources are highlighted in green, blue, and orange for pneumatic controllers, produced water tanks, and oil tanks, respectively. Path names are shown in purple.
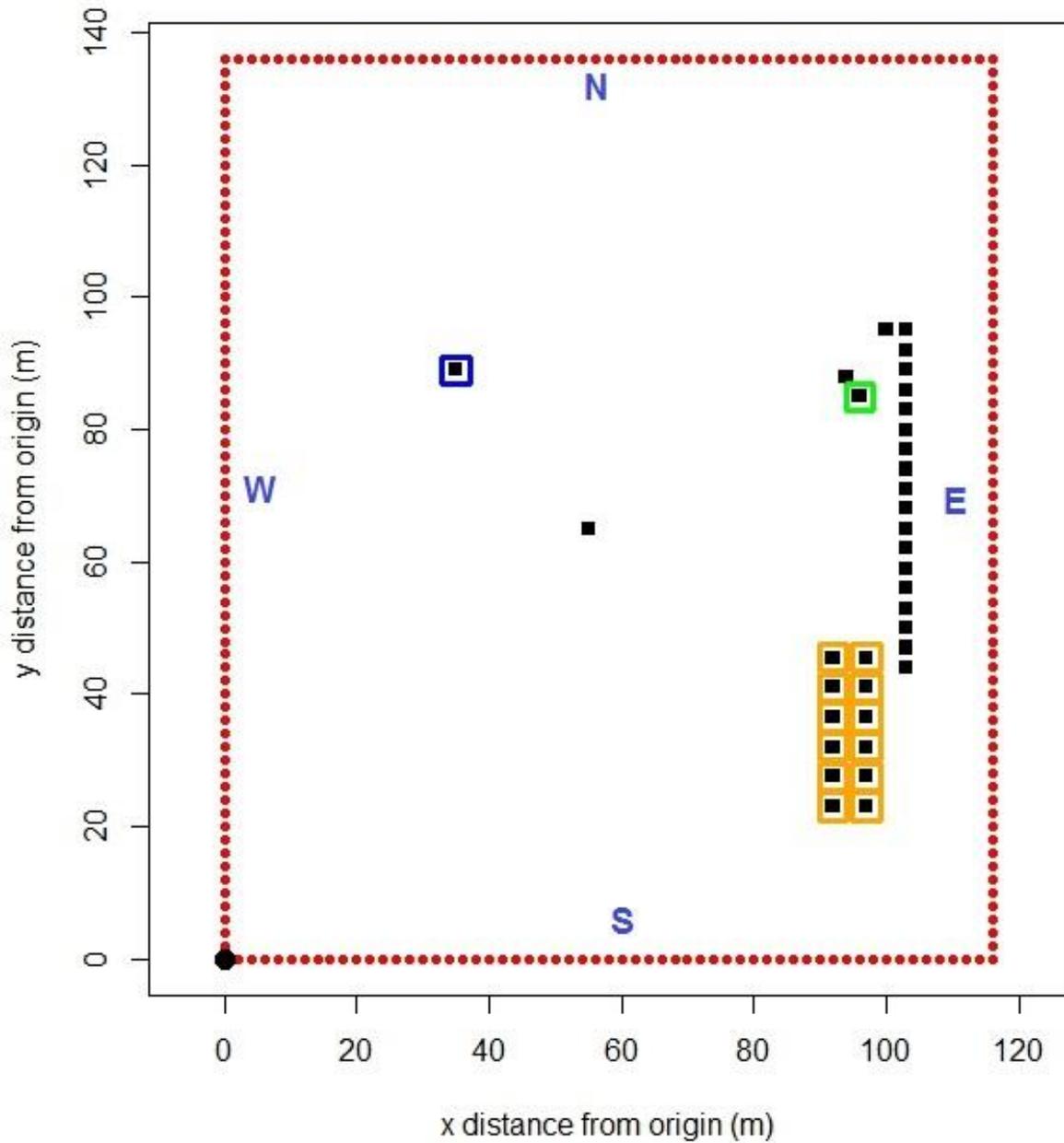
Results & Discussion

Sensitivity Analysis

For the Eagle Ford Shale, wind direction is most commonly from the south-southeast with other directions between 0 - 180^0 also common; winds from directions >180^0 to <360^0 only occur 15% of the time (Figure 7). Therefore, time between detection is highly sensitive to the relative orientation of the sensor and receptor. At 1 m distance between a 150 scfh source and a 0.5 ppm detection limit sensor, average time between detection ranges from 3 hours if the receptor is north-northwest of the source to 30 hours if it is east (Figure 8). The range increases from 4 hours (NNW) to 50 hours (E) at 10 meters, 11 hours (NW) to 290 hours (E) at 100 meters, and 39 hours (NNW) to 1,478 hours (NE) at 200 meters. The effect of orientation is even greater on the 95th percentile highest time between detection (Figure S1); for example, at 10 meters distance the range is from 13 hours (NW) to 269 hours (ENE). At 10 m distance, average time between detection was highly sensitive to relative height differences of the source and receptor (Figure 9). The average detection time of all orientations was 18, 22, 47, and 190 hours at 0, 1, 2, and 3 m relative height difference, respectively; at 4 m difference, emissions were never detected from most orientations. At 100 m distance, average detection time was much less sensitive to relative height difference; the average of all orientations was 86, 90, 97, 106, 133, and 176 hours at 0, 1, 2, 3, 4, and 5 m relative height difference, respectively (Figure S2).

For the Fayetteville and Bakken Shale, wind direction was more variable (Figures S3 and S4). Accordingly, average time between detection was both faster and less affected by orientation than the Eagle Ford. At 100 m distance, average detection time was 37 and 47 hours in the Fayetteville and Bakken, respectively, compared to 86 hours in the Eagle Ford. The orientations with the minimum and maximum time varied by a ratio of 3 and 4 in the Fayetteville and Bakken, respectively, compared to 27 in the Eagle Ford. Full sensitivity results for the Fayetteville and Bakken are shown in Figures S5 – S12.

Figure 7. Eagle Ford Shale meteorological data: Wind rose of Lackland Airforce Base (SFK)
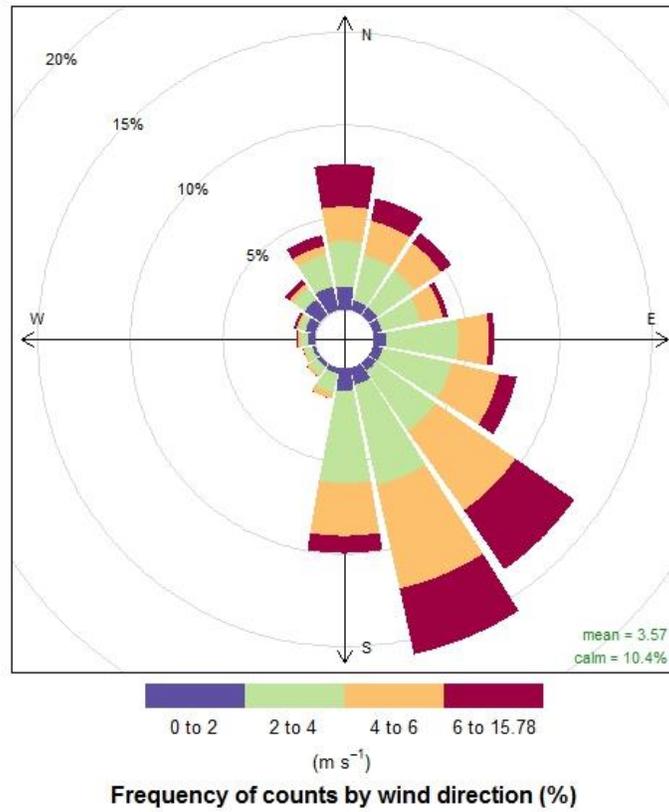
Figure 8. Eagle Ford Shale meteorological data: Sensitivity of average hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and distance of receptor from source
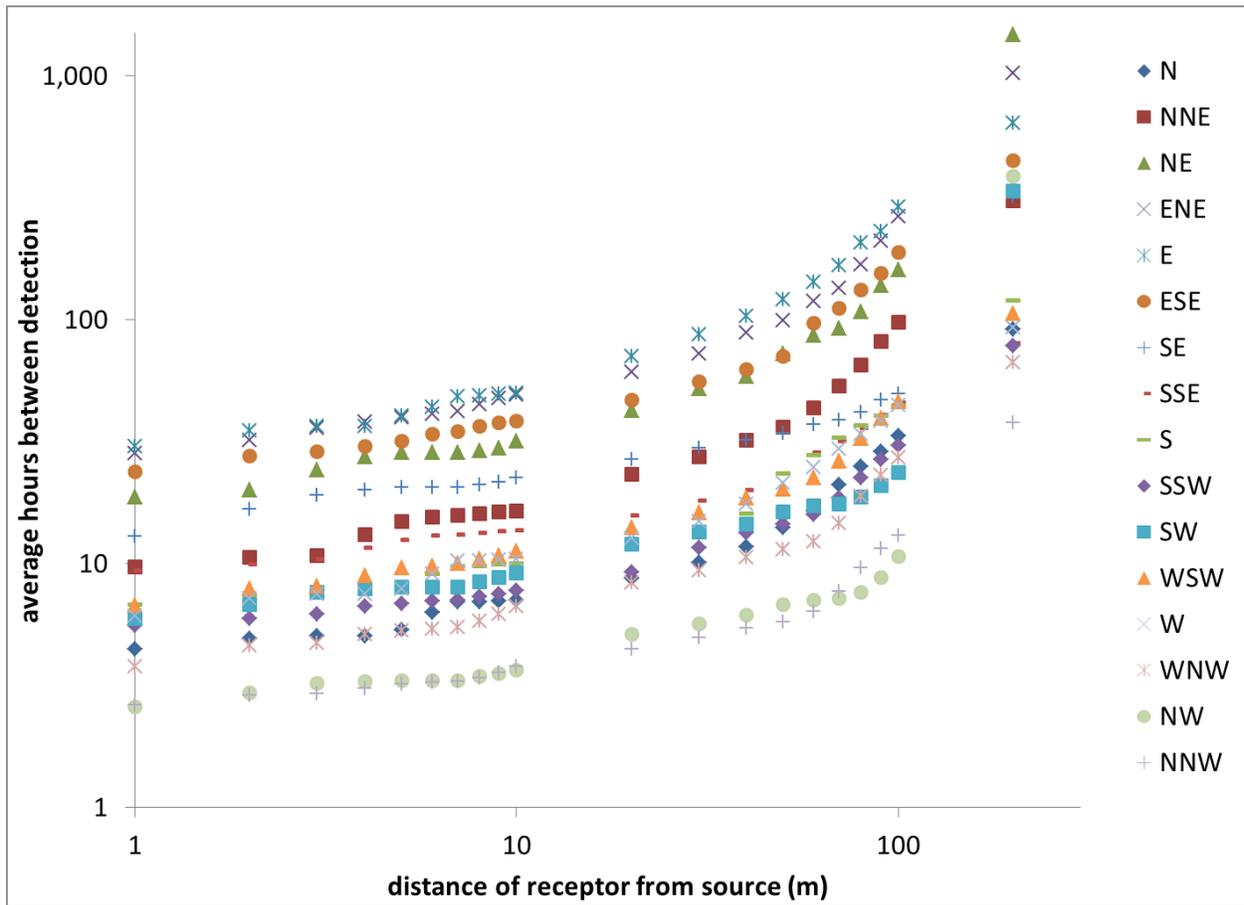
Figure 9. Eagle Ford Shale meteorological data: Sensitivity of average hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and relative height distance between a source and a receptor at 10 m distance
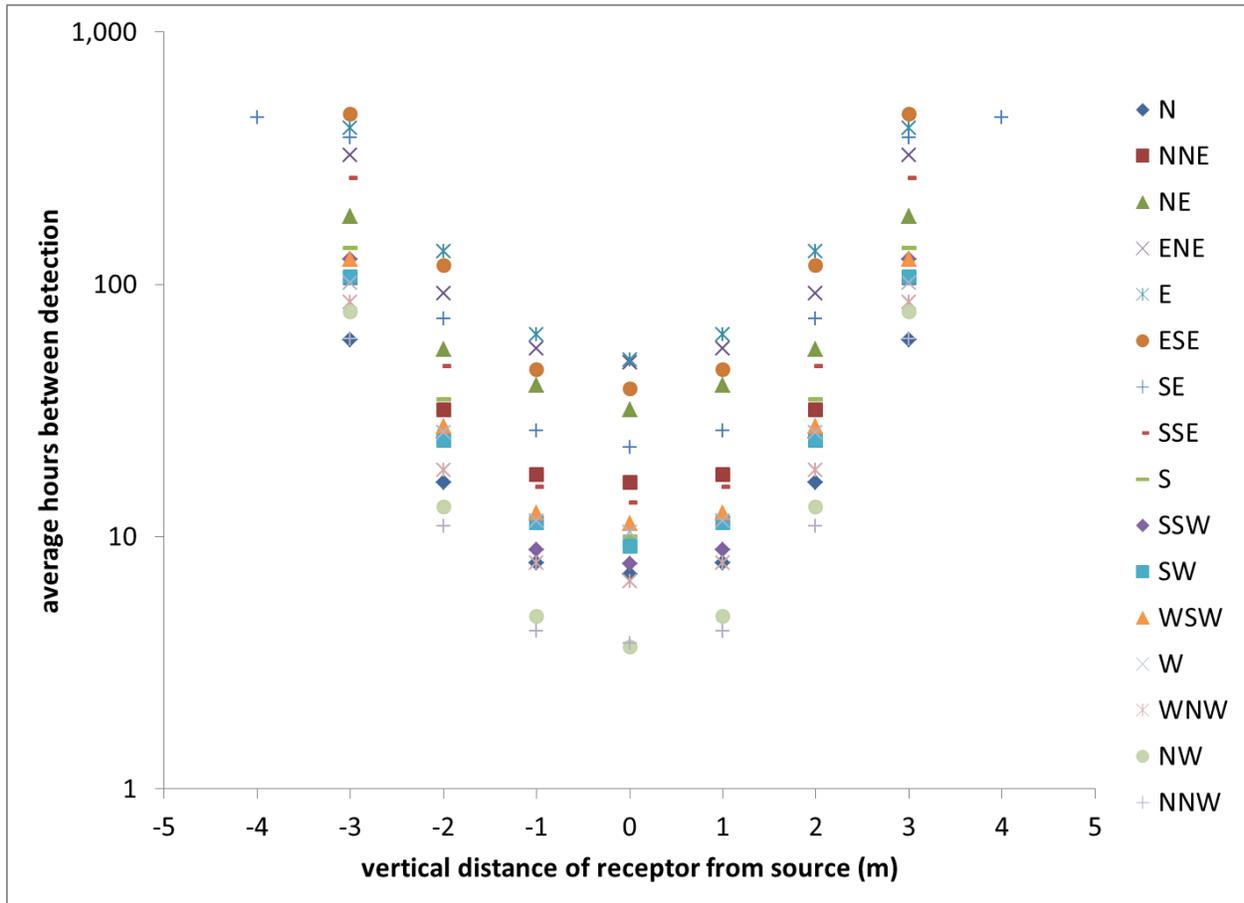
Eagle Ford Shale Well Pad

For the Eagle Ford well pad, average time between detection of a 150 scfh emission source with a 0.5 ppm detection limit point sensor ranges from 3 to 550 hours for every source and receptor combination (Figure S13). Site-level performance, defined as the median of all sources' detection time, ranges from 12 to 221 hours with the best receptors located on paths W and I1 near the origin (Figure 10; the site is offset 67^0 from N so the y-axis is oriented ENE). The receptor with the lowest median source detection time, R366, has a site-level performance of 12 average hours between detection (95th percentile = 58 hours). At this receptor, a sensor can detect emissions most quickly from source C67, a connector near path S, with an average detection time of 5 hours, and least quickly from S2C, a 3 m high separator leak point at the opposite side of the site, with an average detection time of 48 hours (Figure 11). Receptor R366 is close to equipment and therefore may not be suitable if a sensor is not intrinsically safe; receptor R282 on the fence line path W performs nearly as well at a safe distance from the nearest equipment.

For an open-path sensor, average time between detection of a 150 scfh leak with a 0.5 ppm detection limit sensor ranges from 2 to 200 hours for every source and path combination (Figure S14). The path with the best site-level performance (E) has an average time of 5 hours between detection (Figure 12). For individual sources, performance at path E is roughly opposite as receptor R366: C26, a connector near path E, has the quickest time to detection (3 hours), and C72, a connector near path S, has the longest time (23 hours). The good performance of path E seems counterintuitive since individual receptors along the path performed poorly, but can be explained by the relative orientation of the majority of sources. Due to the site offset, winds from the NW to NE blow emissions from the separators and most connectors less than 20 m downwind to path E, which results in high path-average concentration. In contrast, during more common SE winds, most sources are more than 50 m downwind of path I2 and therefore the path-average concentration is less likely to be above the detection limit.

For a 6 scfh emissions from equipment leaks, performance is poor with a 0.5 ppm detection limit sensor (Figure S15). The receptor with the best site-level performance, R82, has a median source detection time of 157 hours and shortest time of 48 hours for source C47, but never detects emissions from source S2C (Figure S16). The best path (E) is even worse with a median source average time of over 6,000 hours

and inability to detect emissions from source S2C. None of the evaluated receptor points or paths at this site was able to detect 6 scfh leaks with a 0.5 ppm sensor from all sources. However, if a sensor has a 0.1 ppm detection limit, then performance detecting 6 scfh leaks is greatly improved (Figure S17). A point sensor at receptor R82 has a median source average detection time of 66 hours with the longest time of 421 hours for source S2C (Figure 13). An open-path sensor with a 0.1 ppm detection limit has better site-level performance than a point sensor with a source median of 24 hours, but requires about 200 hours longer for the worst source (Figure 14).

Table 1. Eagle Ford well pad: Average and 95th percentile hours between detection of either a 150 scfh emission rate from all sources or a 6 scfh emission rate from equipment leak sources. Values are shown for three sensor detection limits of resolution of enhancement from background. Site-level best receptor or path locations and hours between detection are determined by the median of individual sources' time between detection. ND = no detection during 10 years of meteorological data.

		average and 95th percentile hours between detection for three sensor detection limits					
		0.5 ppm		0.1 ppm		1 ppm	
		mean	95th	mean	95th	mean	95th
150 scfh all sources	best receptor for all sources (R366)	12	58	8	32	18	88
	best source for receptor R366 (C67)	5	21	5	21	6	23
	worst source for receptor R366 (S2C)	48	210	18	92	78	296
	best fence line receptor for all sources (R282)	12	59	6	25	22	101
	best source for receptor R282 (C35)	7	25	5	20	24	110
	worst source for receptor R282 (S2C)	39	169	12	59	142	534
	best path for all sources (E)	5	18	4	13	9	37
	best source for path E (C26)	3	11	3	10	4	16
	worst source for path E (C72)	23	116	13	59	29	148
6 scfh leaks only	best receptor for all sources (R82)	157	628	66	277	526	1,331
	best source for receptor R82 (C47)	48	213	25	126	99	443
	worst source for receptor R82 (S2C)	ND	ND	421	1,765	ND	ND
	best path for all sources (E)	6,085	16,486	24	118	ND	ND
	best source for path E (C12)	254	573	10	44	1,247	2,675
	worst source for path E (S2C)	ND	ND	627	1,523	ND	ND

Figure 10. Eagle Ford well pad: Site-level average hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit point sensor for every receptor. Site-level metrics are based on the median of individual sources' average hours. Colors show the quantile of hours between detection (1st, 2.5th, 5th, 10th, 25th, 50th, & 75th percentiles).
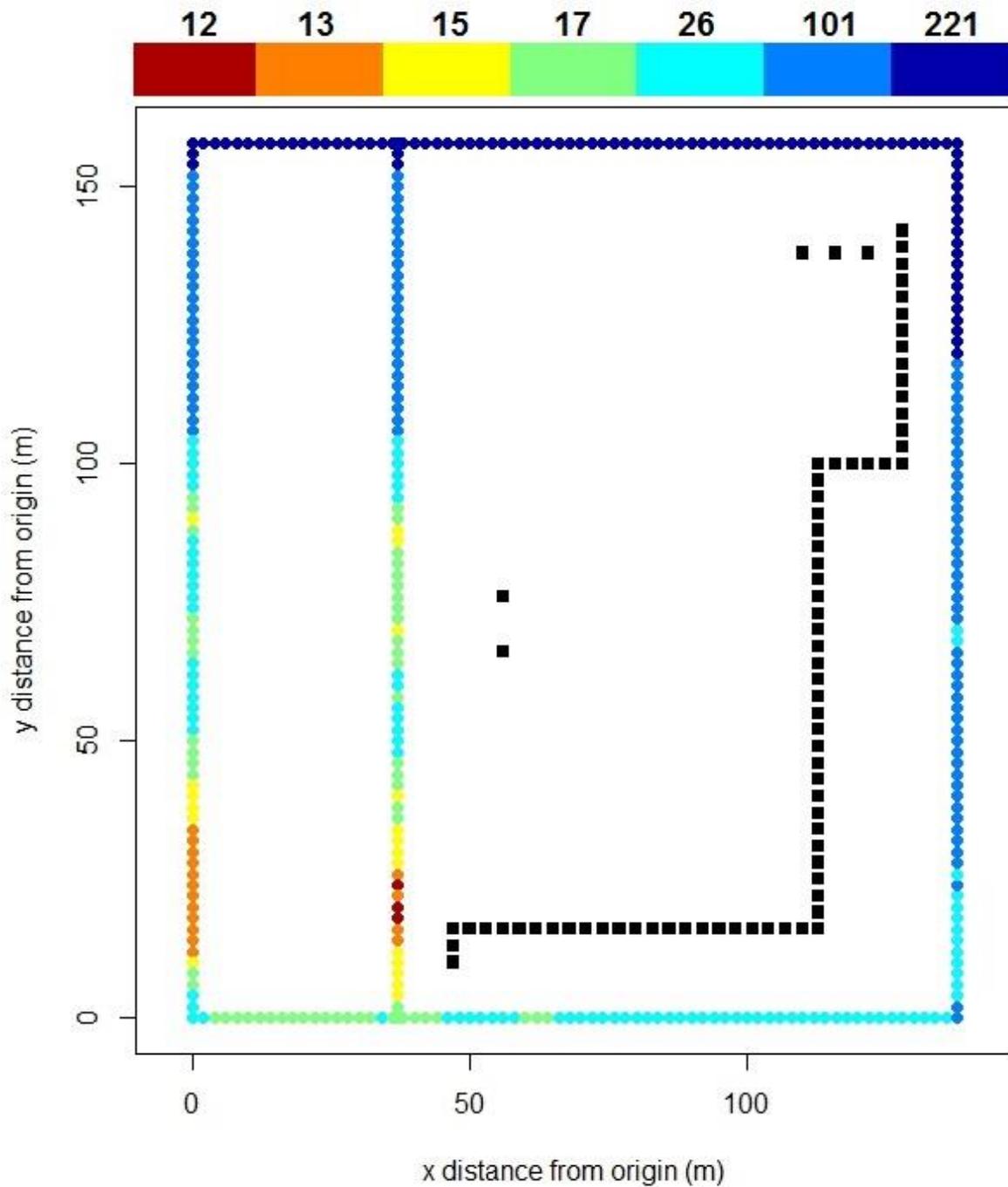
Figure 11. Eagle Ford well pad: Average hours between detection of a 150 scfh emission rate from individual sources with a 0.5 ppm detection limit point sensor located at the best receptor (R366, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
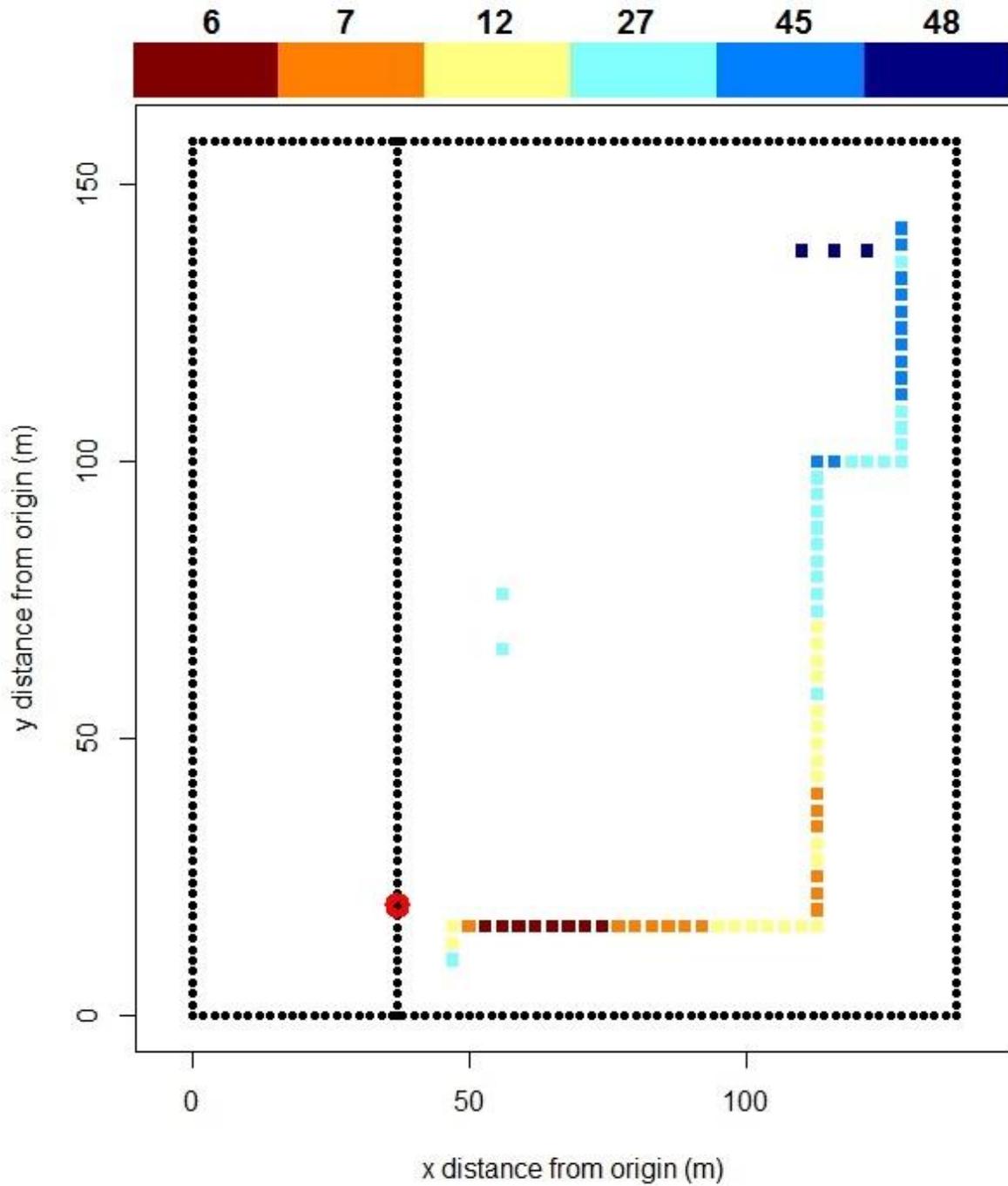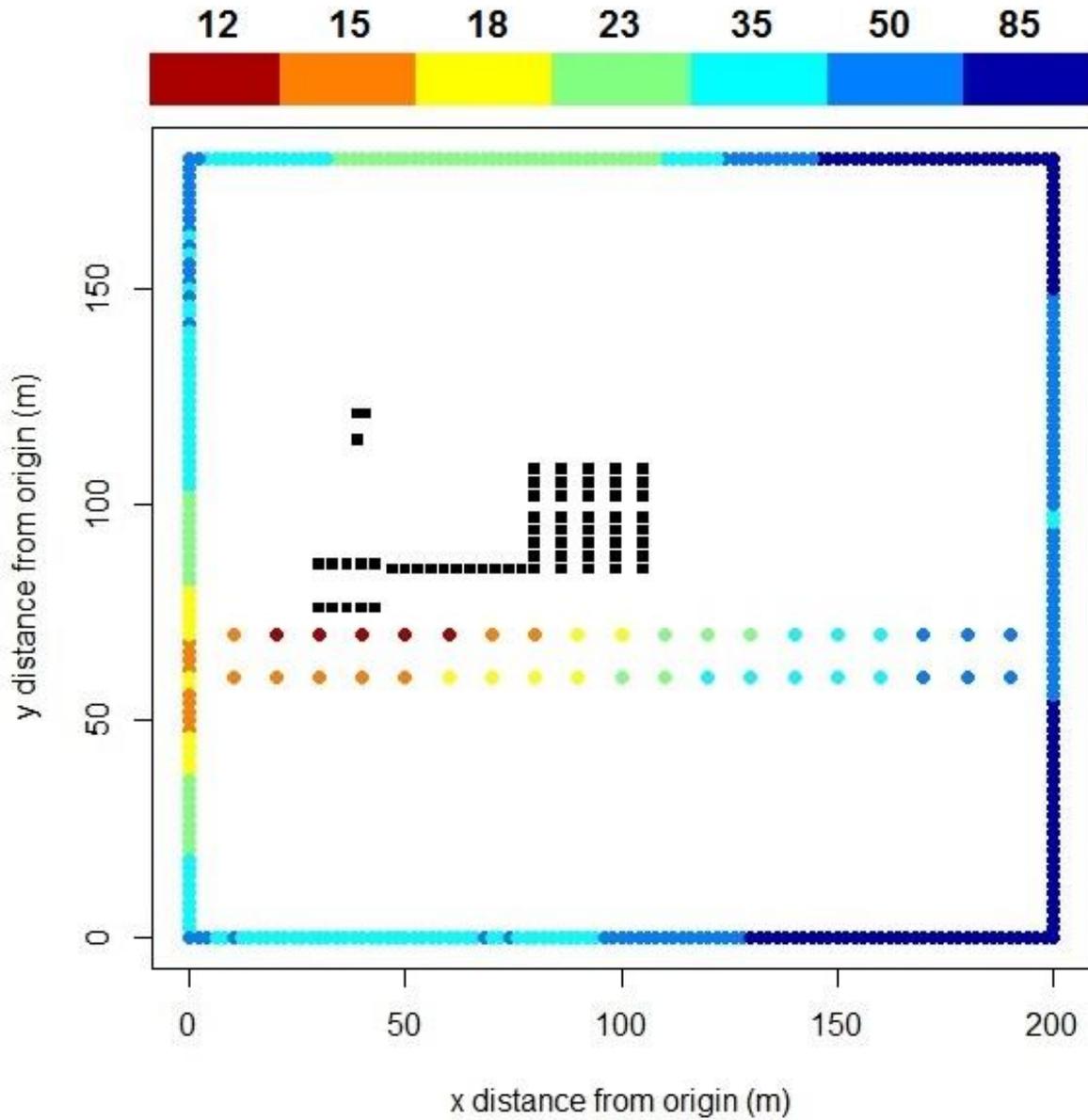
Figure 12. Eagle Ford well pad: Average hours between detection of a 150 scfh emission rate from individual sources with a 0.5 ppm detection limit open path sensor deployed across the best path (E1, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
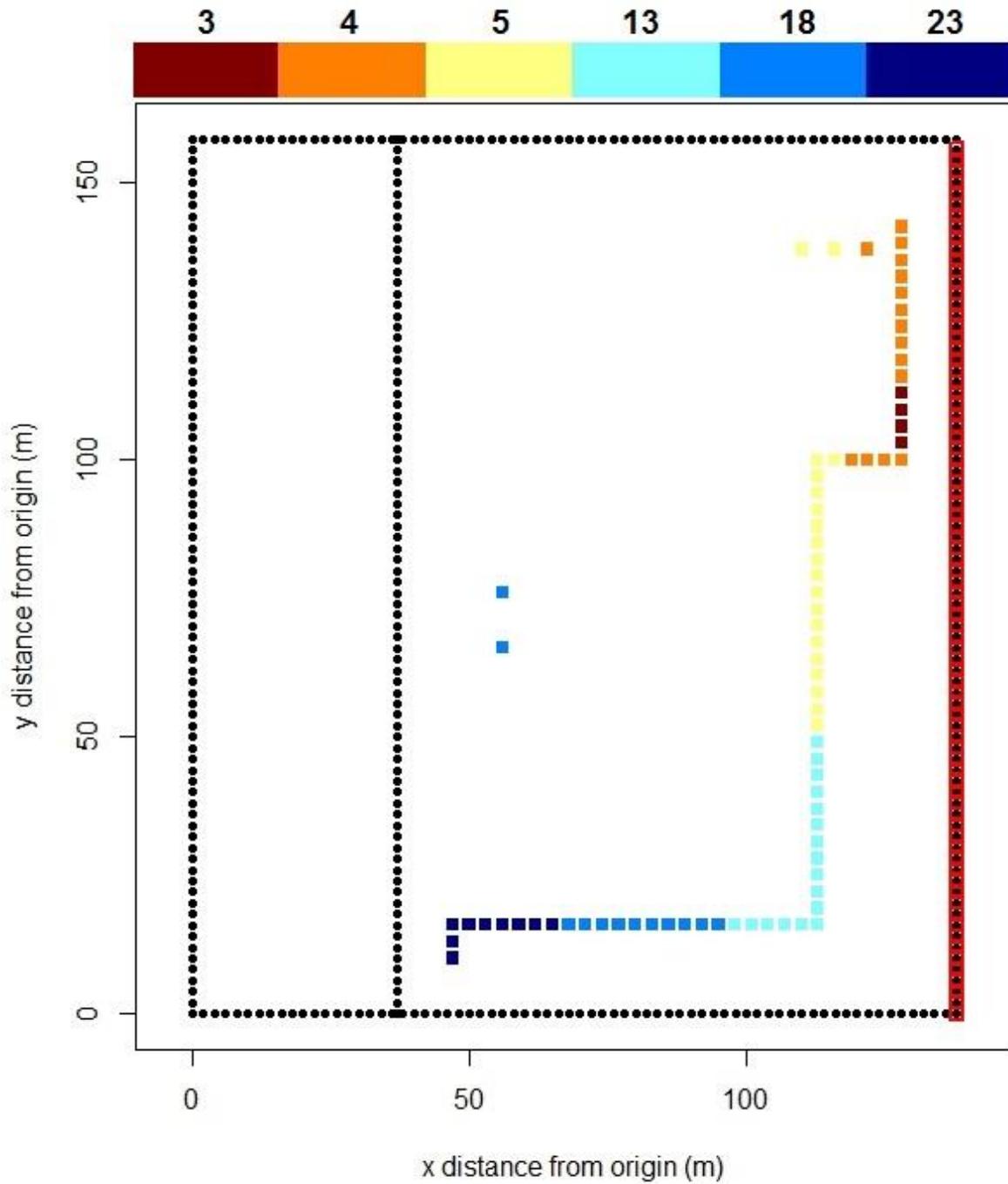
Figure 13. Eagle Ford well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.1 ppm detection limit point sensor at the best receptor (R82, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
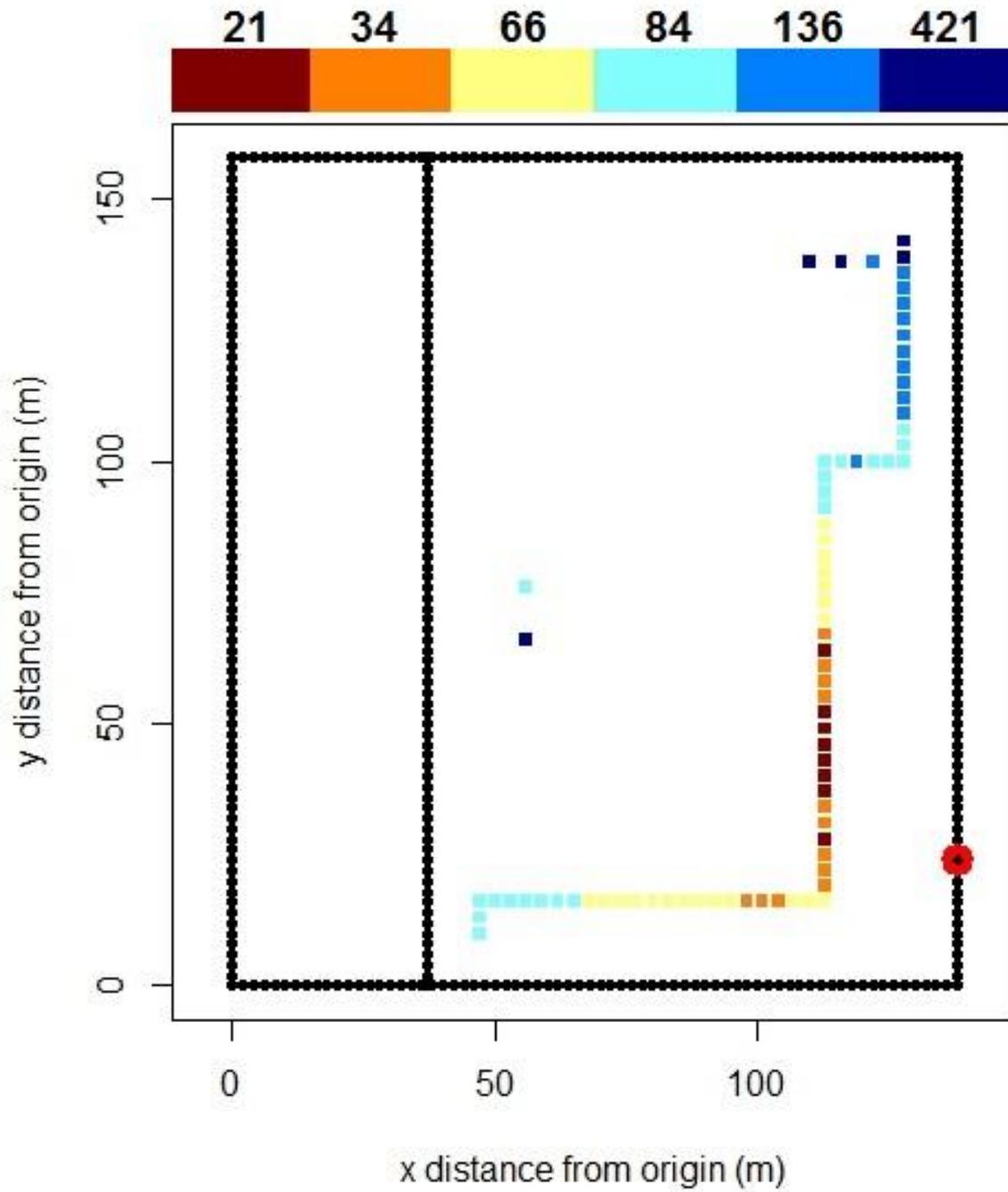
Figure 14. Eagle Ford well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.1 ppm detection limit open path sensor deployed across the best path (E1, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
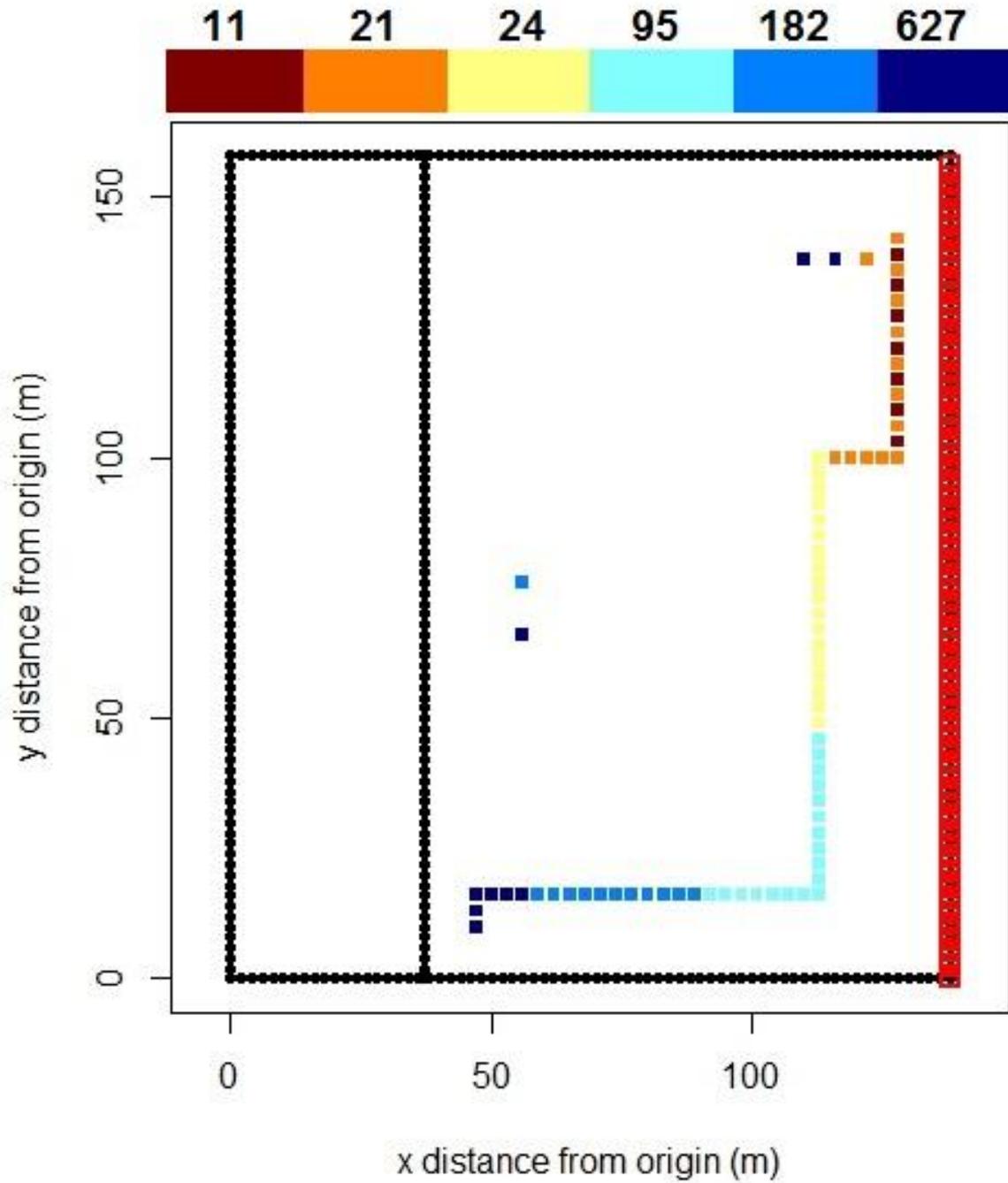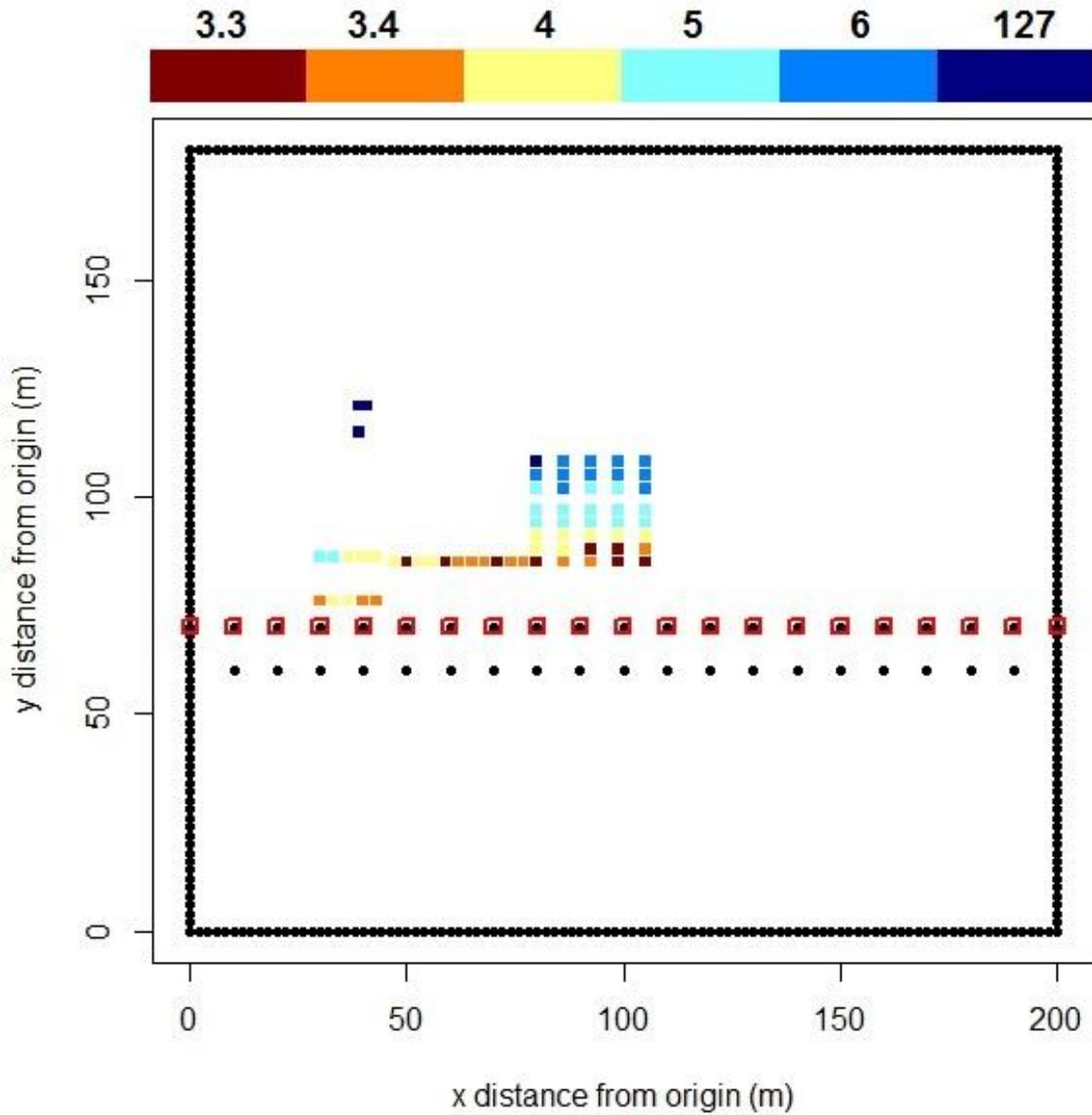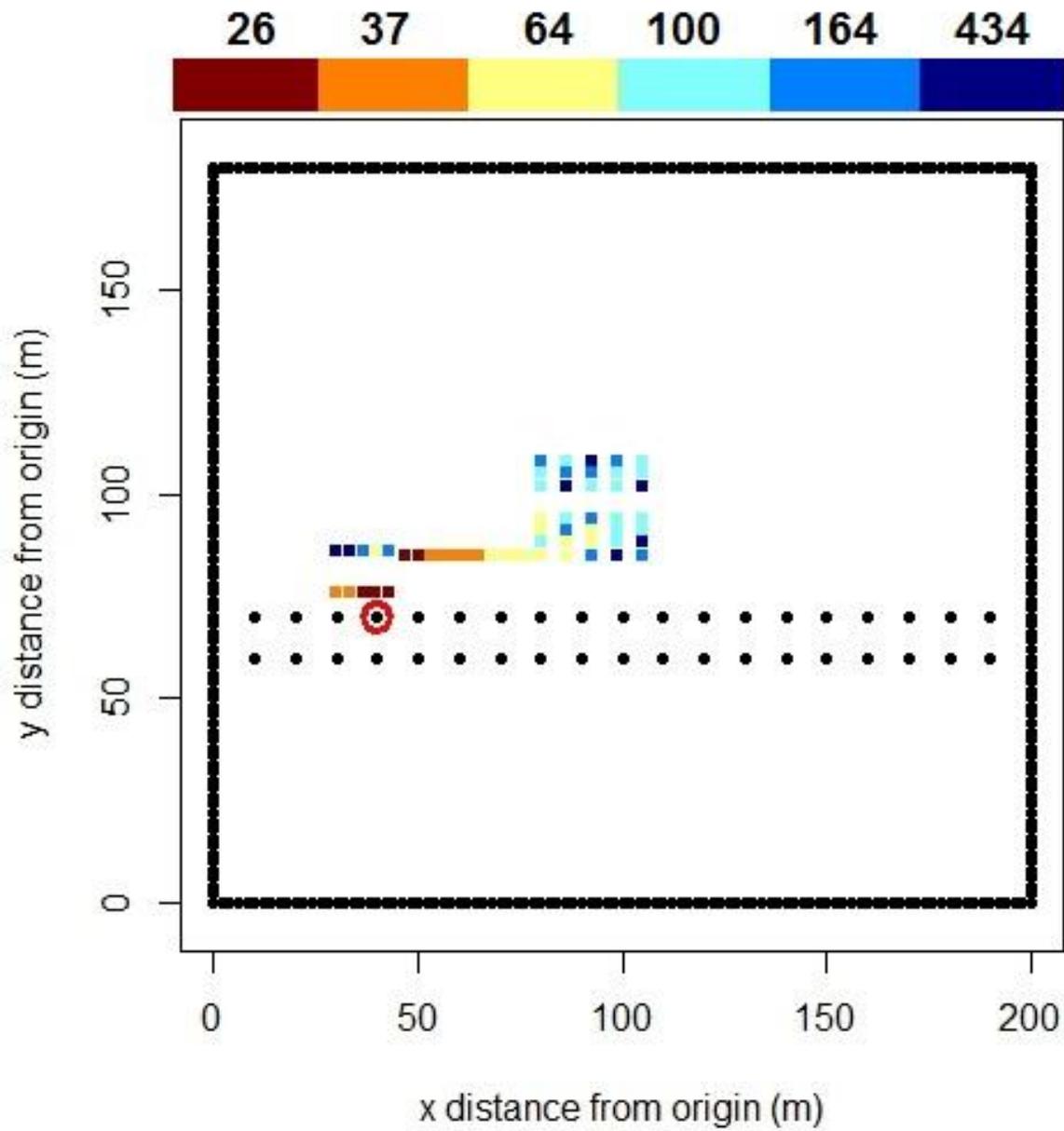
Fayetteville Shale Well Pad

For the Fayetteville site, average time between detection of a 150 scfh emission source with a 0.5 ppm detection limit point sensor ranges from 8 to 401 hours for every source and receptor combination (Figure S19). Site-level detection time is quickest at receptors located on paths I1, I2, and W southwest of most sources (Figure 15). The best location for a point sensor is R432 on path I1, which has a median source average detection time of 12 hours and a longest time of 28 hours for tank hatch source TH3 (Figure 16). An alternative location for a non-intrinsically safe sensor, R355 on the west fence line, has a slightly higher detection time of 14 hours (Figure S20). For an open-path sensor with a 0.5 ppm detection limit, average detection time ranged from 3 to 346 hours for every source and path combination (Figure S21). Path I1 has the quick median source detection time of 4 hours but requires over 4 times as long as a point sensor at R432 to detect emissions from the worst source TH3.

Compared to the Eagle Ford site, a 0.5 ppm detection limit sensor has better performance detecting 6 scfh leaks at the Fayetteville site. There are several receptors on the west side of paths I1 and I2 where a point sensor can detect emissions with an average time of less than 100 hours (Figure S22). The best point receptor, R433, has a median source detection time of 64 hours with the longest time of 434 hours for source C25 (Figure 18). An alternative point R344 along the west fence line that is suitable for a non-intrinsically safe sensor has good site-level performance but very rarely can detect emissions from the worse source C26 (Figure S23). A 0.5 ppm open path sensor deployed across the best path, I1, has poorer performance than the point sensors with a median source detection time of 227 hours (Figure S24). Sensors with 0.1 ppm detection limits have excellent performance detecting 6 scfh equipment leaks (Figure S25). The best location, R433, has a median source detection time of 21 hours and only 29 hours for the worst source (Figure S26) with a 0.1 ppm point sensor; alternative location R344 has poorer performance but still can detect all sources within 100 hours (Figure S27). A 0.1 ppm detection limit open path monitor deployed across path I1 can detect 6 scfh leaks within a median source time of 14 hours and a longest source time of 41 hours (Figure S28).

Table 2. Fayetteville well pad: Average and 95th percentile hours between detection of either a 150 scfh emission rate from all sources or a 6 scfh emission rate from equipment leak sources. Values are shown for three sensor detection limits of resolution of enhancement from background. Site-level best receptor or path locations and hours between detection are determined by the median of individual sources' time between detection. ND = no detection during 10 years of meteorological data.

		average and 95th percentile hours between detection for three sensor detection limits					
		0.5 ppm		0.1 ppm		1 ppm	
		mean	95th	mean	95th	mean	95th
150 scfh all sources	best receptor for all sources (R432)	12	67	10	51	14	81
	best source for receptor R432 (M3)	8	43	7	40	8	45
	worst source for receptor R432 (TH3)	28	136	19	95	46	199
	best fence line receptor for all sources (R355)	14	81	10	58	21	114
	best source for receptor R355 (M2)	10	56	9	48	11	63
	worst source for receptor R355 (TH2)	37	169	16	88	91	434
	best path for all sources (I1)	4	16	3	11	7	27
	best source for path I1 (C26)	3	12	3	10	5	19
	worst source for path I1 (TH3)	127	509	9	45	632	2,039
6 scfh leaks only	best receptor for all sources (R433)	64	289	21	113	102	445
	best source for receptor R433 (M5)	15	78	11	55	21	111
	worst source for receptor R433 (C25)	434	1,535	29	145	12,408	21,306
	best fence line receptor for all sources (R344)	93	415	49	242	224	677
	best source for receptor R344 (S1A)	42	211	14	78	60	286
	worst source for receptor R344 (C10)	22,404	33,022	92	404	ND	ND
	best path for all sources (I1)	227	781	14	67	2,822	5,939
	best source for path I1 (M1)	57	237	8	33	153	579
	worst source for path I1 (W3)	11,073	37,387	41	176	ND	ND

Figure 15. Fayetteville well pad: Site-level average hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit point sensor for every receptor. Site-level metrics are based on the median of individual sources' average hours. Colors show the quantiles of time between detection (1st, 2.5[th], 5th, 10th, 25th, 50th, & 75th percentiles).

Figure 16. Fayetteville well pad: Average hours between detection of a 150 scfh emission rate from individual sources with a 0.5 ppm detection limit point sensor located at the best receptor (R432, highlighted in red) Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
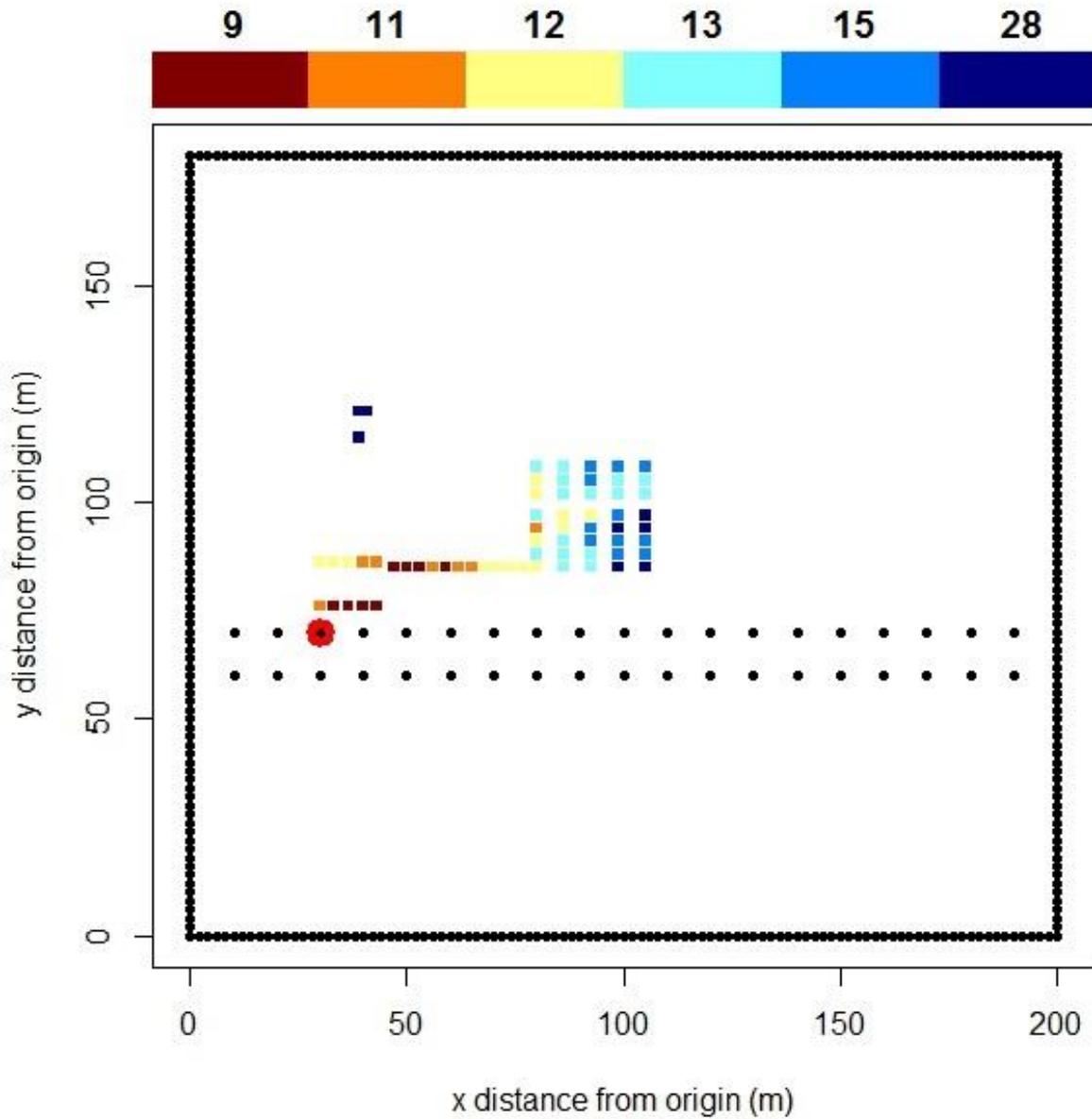
Figure 17. Fayetteville well pad: Average hours between detection of a 150 scfh emission rate from individual sources with a 0.5 ppm detection limit open path sensor deployed across the best path (I1, highlighted in red) Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).

Figure 18. Fayetteville well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.5 ppm detection limit point sensor located at the best receptor (R433, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).

Bakken Shale Well Pad

For the Bakken site, average time between detection of a 150 scfh emission source with a 0.5 ppm detection limit point sensor ranges from 9 to 1,490 hours for every source and receptor combination (Figure S29). Since the Bakken site has a high diversity of source heights including twelve tanks modeled at 6 m height, best point and path locations were compared for receptors at 1 m and 5 m height. At 1 m height, the best receptors for point sensors are in the central section of the east fence line (Figure 19), while at 5 m height the best receptors are at the northern section of the east fence line and eastern section of the south fence line (Figure S30). The best individual receptor, R1M96, is at 1 m height and has a median source average detection of 11 hours and a worse source detection time of 219 hours for source SEP1C, a 6 m high separator leak (Figure 20). In contrast, the best receptor at 5 m height (R5M50), has a longer median source average detection of 20 hours but only 48 hours for the worst source (Figure 21). Therefore, higher sensor deployment is more appropriate for this site since it better covers all sources with only minor reduction in median detection time. For an open-path sensor with a 0.5 ppm detection limit, average time between detection ranges from 2 to 8,467 hours for every source and path combination (Figure S31). Similar to point sensors, the height of the path involves a tradeoff of the median source and worst source detection time. An open-path sensor deployed along the eastern fence line at 1 m height (E1) has a median source detection time of 3 hours with an average time of 8,467 hours for the worst source (Figure 22), while a sensor deployed at 5 m height (E5) has a median source and worst source detection of 48 and 298 hours, respectively (Figure S32).

For equipment leaks, receptor locations at 1 m height are most appropriate since the vast majority of leak sources are at 1 m height. The receptor with the lowest median source detection time of a 6 scfh leak (R1M100) has good detection times for point sensors with either a 0.5 ppm detection limit (44 hours) or 0.1 ppm detection limit (14 hours), but never detects emissions from source SEP1C (Figure S33). There are no individual receptors that a 0.5 ppm sensor can detect 6 scfh emissions from all leak sources, but for a 0.1 ppm detection limit receptor R1M251 has the shortest worst source detection time of 845 hours; the source median time for this receptor is 190 hours (Figure 23). No evaluated paths can detect 6 scfh leaks with a 0.5 ppm detection limit open path sensor. For a 0.1 ppm detection limit open path sensor,

the eastern fence line (E1) has a median source detection time of 11 hours but emissions are never detected from source SEP1C (Figure S34).

Due to the large number of tanks on this site, emissions were also evaluated from tank venting to determine how often enhancement from normal venting emissions would exceed 0.1 ppm. Based on an uncontrolled tank venting emission factor of 83 scfh, emissions would cause enhancement ≥0.1 ppm every 16 – 21 hours at R1M96 (Figure 24), 36 – 178 hours at R5M50 (Figure S32), and 15 – 401 hours at path E (Figure S33). These estimates are based on average emission rates, while in reality tank flashing emissions may be intermittent with higher rates during 30 – 120 second periods following separator dumps. Since tank flashing emissions regularly will cause enhancements ≥0.1 ppm, a system designed to detect equipment leaks based on 0.1 ppm enhancement will have false alarms if it cannot distinguish vented emissions. One solution to this issue is for sensors employ algorithms that evaluate the concentration time series. If concentration increases periodically during steady winds, then it is likely from an intermittent source such as tank flashing rather than a fugitive leak with a steady emission rate.

Table 3. Bakken well pad: Average and 95th percentile hours between detection of either a 150 scfh emission rate from all sources or a 6 scfh emission rate from equipment leak sources. Values are shown for three sensor detection limits of resolution of enhancement from background. Site-level best receptor or path locations and hours between detection are determined by the median of individual sources' time between detection. ND = no detection during 10 years of meteorological data.

		average and 95th percentile hours between detection for three sensor detection limits					
		0.5 ppm		0.1 ppm		1 ppm	
		mean	95th	mean	95th	mean	95th
150 scfh all sources	best receptor at 1 m height for all sources (R1M96)	11	62	10	53	13	73
	best source for receptor R1M96 (C6)	9	51	8	43	12	65
	worst source for receptor R1M96 (SEP1C)	219	966	57	247	2,240	7,490
	best receptor at 5 m height for all sources (R5M50)	20	100	13	70	48	198
	best source for receptor R5M50 (TH7)	11	59	9	51	13	69
	worst source for receptor R5M50 (PWTH1)	48	197	16	82	367	1,197
	best path at 1 m height for all sources (E1)	3	13	2	7	9	39
	best source for path E1 (C7)	2	7	2	6	4	15
	worst source for path E1 (PWTH1)	8,467	27,733	8	44	ND	ND
	best path at 5 m height for all sources (E5)	48	187	6	27	769	3,161
	best source for path E5 (TH12)	5	20	2	7	12	57
	worst source for path E5 (PJ1)	298	1,530	5	22	ND	ND
6 scfh leaks only	best receptor for all sources (R1M100)	44	200	14	73	80	323
	best source for R1M100 (C11)	33	153	12	60	62	281
	worst source for R1M100 (SEP1C)	ND	ND	ND	ND	ND	ND
	receptor with least worst source for all sources (R1M251)	ND	ND	190	727	ND	ND
	best source for R1M251 (PJ1)	222	748	92	386	ND	ND
	worst source for R1M251 (C7)	ND	ND	845	2,039	ND	ND
	best path for all sources (E1)	ND	ND	11	47	ND	ND
	best source for path E1 (C12)	ND	ND	10	45	ND	ND
	worst source for path E1 (SEP1B, SEP1C)	ND	ND	ND	ND	ND	ND

Figure 19. Bakken well pad: Site-level average hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit point sensor for every receptor. Site-level metrics are based on the median of individual sources' average hours. Colors show the quantile of hours between detection (1st, 2.5th, 5th, 10th, 25th, 50th, & 75th percentiles).
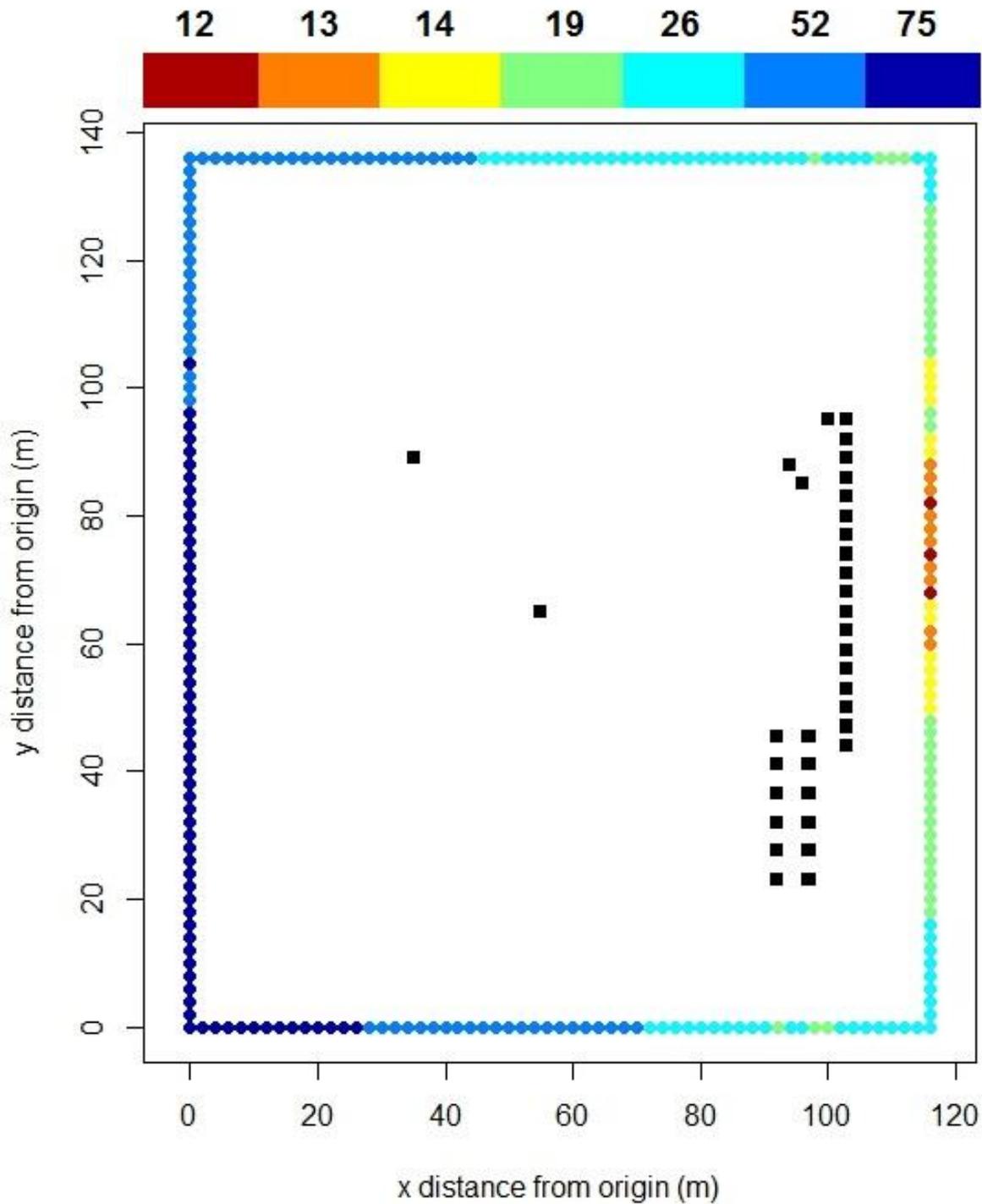
Figure 20. Bakken well pad: Average hours between detection of a 150 scfh emission rate from individual sources with a 0.5 ppm detection limit point sensor located at the best receptor at 1 m height (R1M96, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
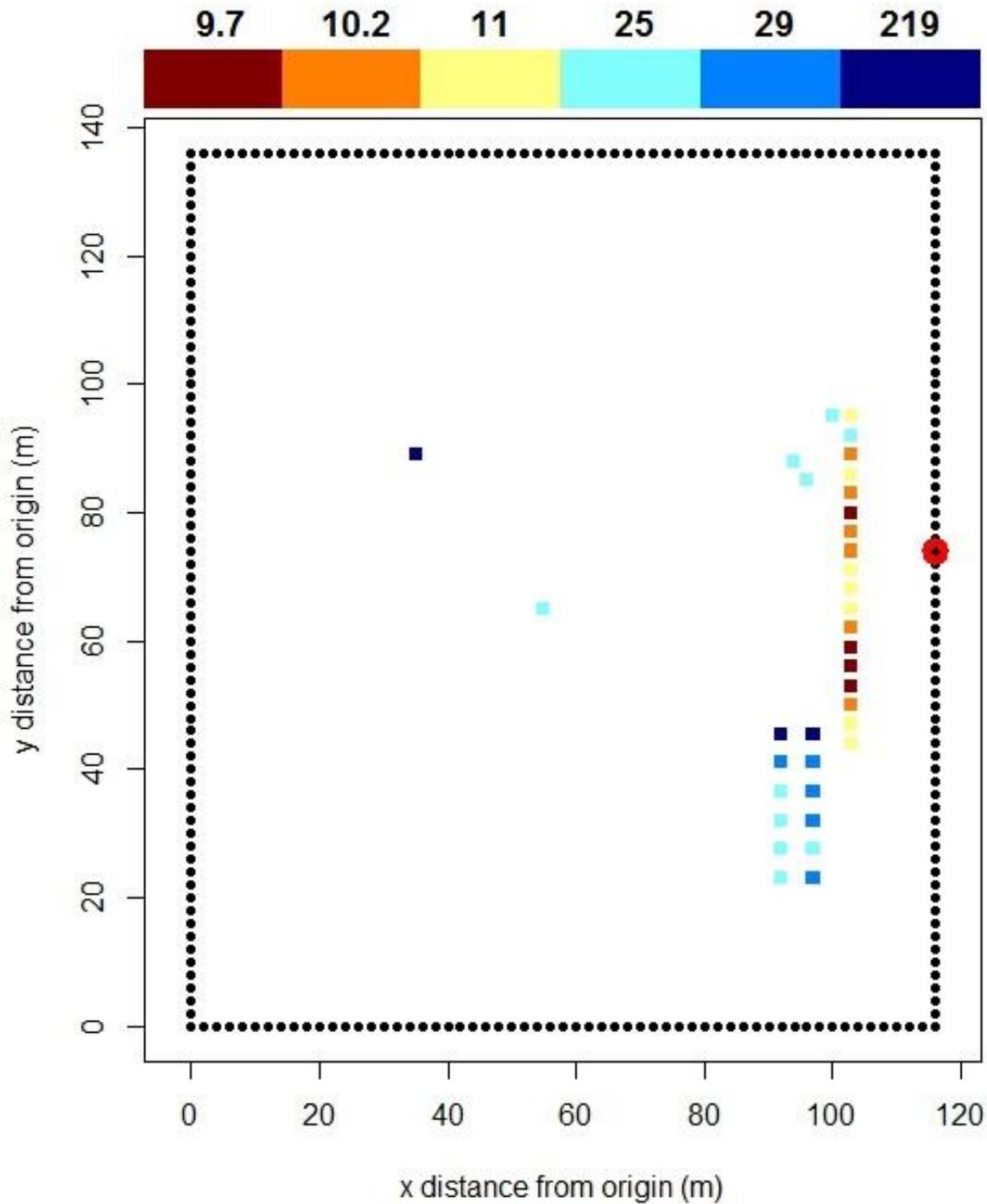
Figure 21. Bakken well pad: Average hours between detection of a 150 scfh emission rate from individual sources with a 0.5 ppm detection limit point sensor located at the best receptor at 5 m height (R5M50, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
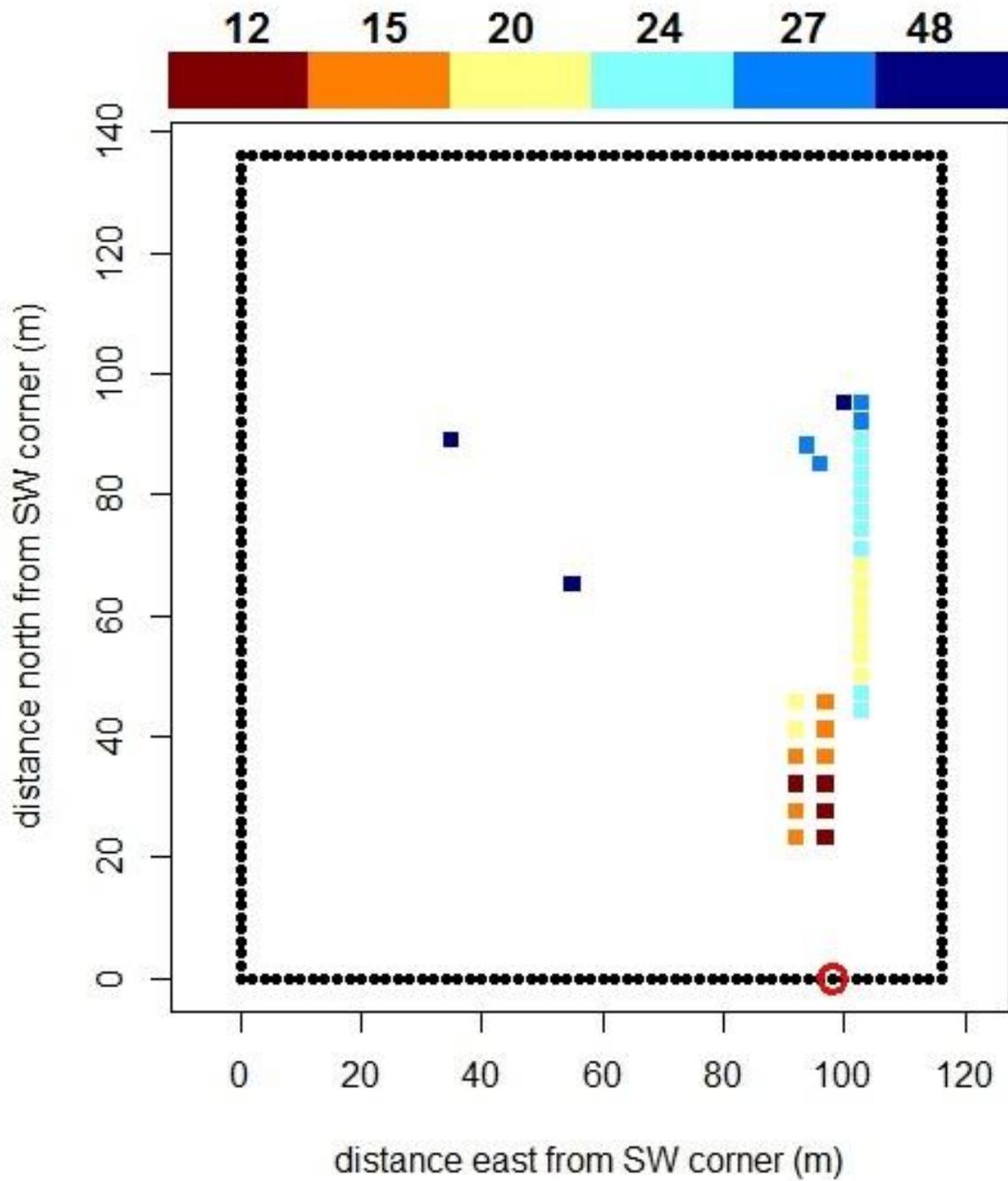
Figure 292. Bakken well pad: Average hours between detection of a 150 scfh emission rate from individual sources with a 0.5 ppm detection limit open path sensor deployed across the best path at 1 m (E1, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
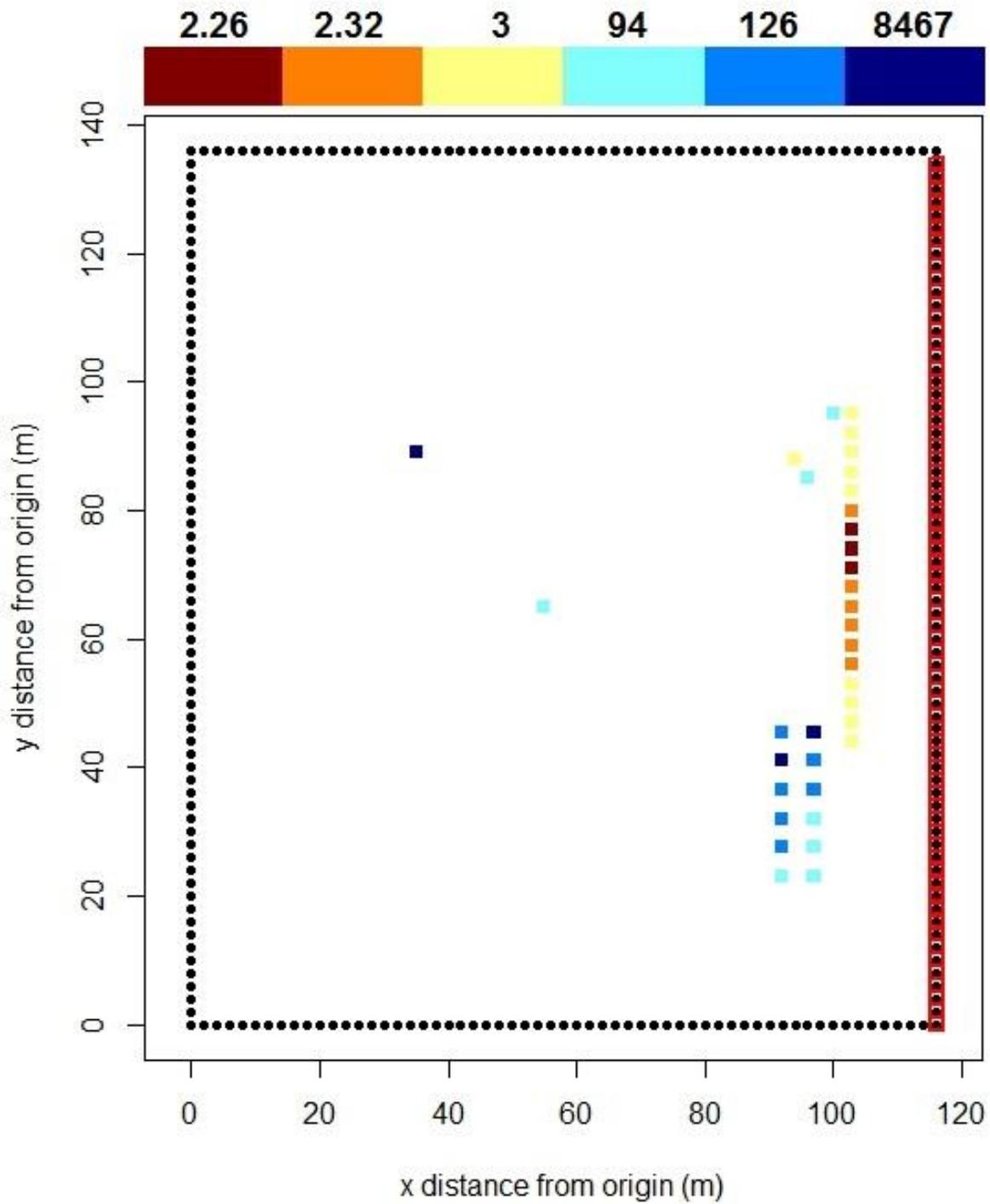
Figure 23. Bakken well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.1 ppm detection limit point sensor located at the receptor with the lowest average hours between detection for the equipment leak source with the highest time between detection. Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
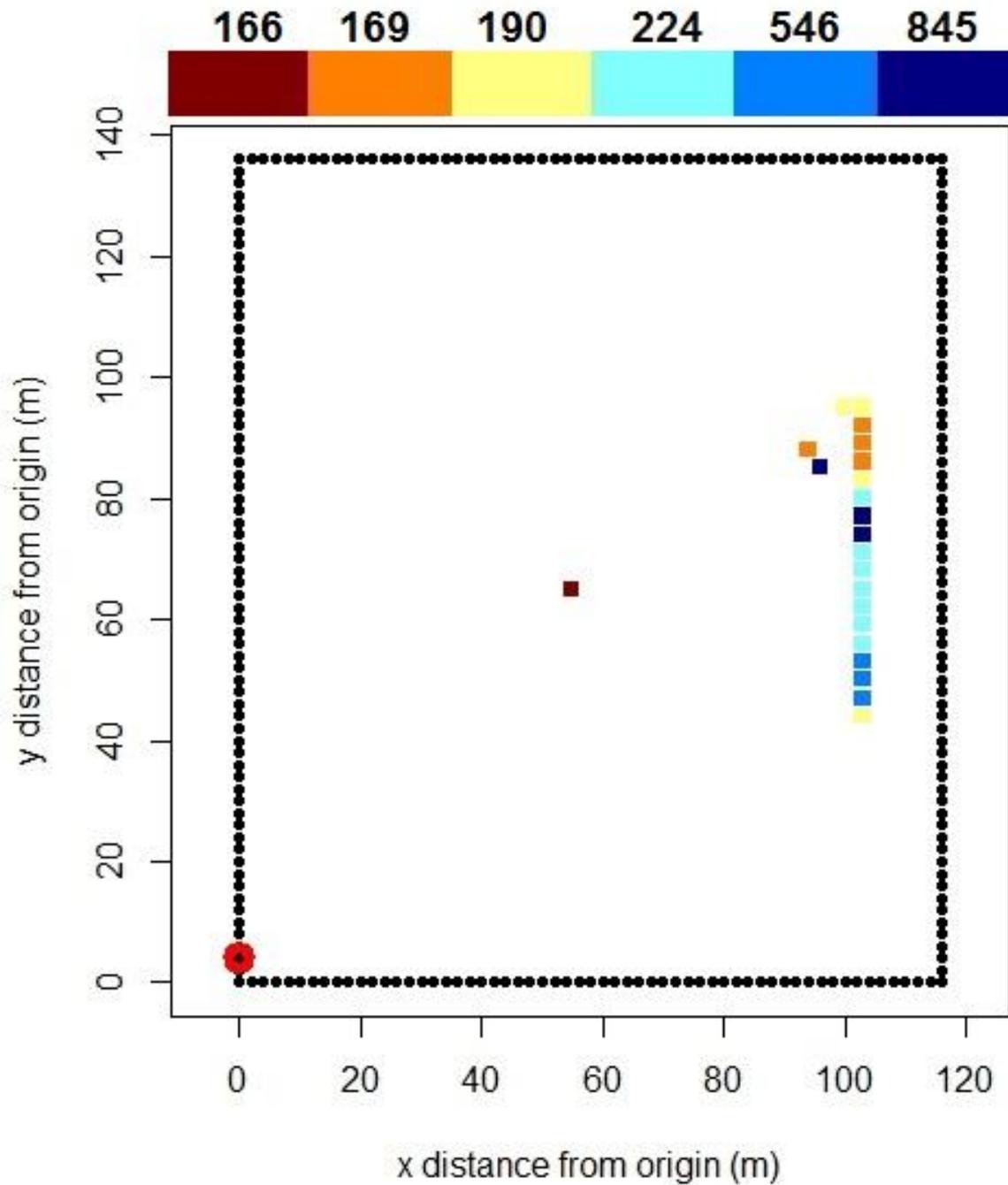
Figure 24. Bakken well pad: Average hours between enhancement ≥ 0.1 ppm from normal emissions of individual vented sources with a sensor located at the receptor with the lowest median of individual equipment leak sources' average hours between detection (R1M100, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
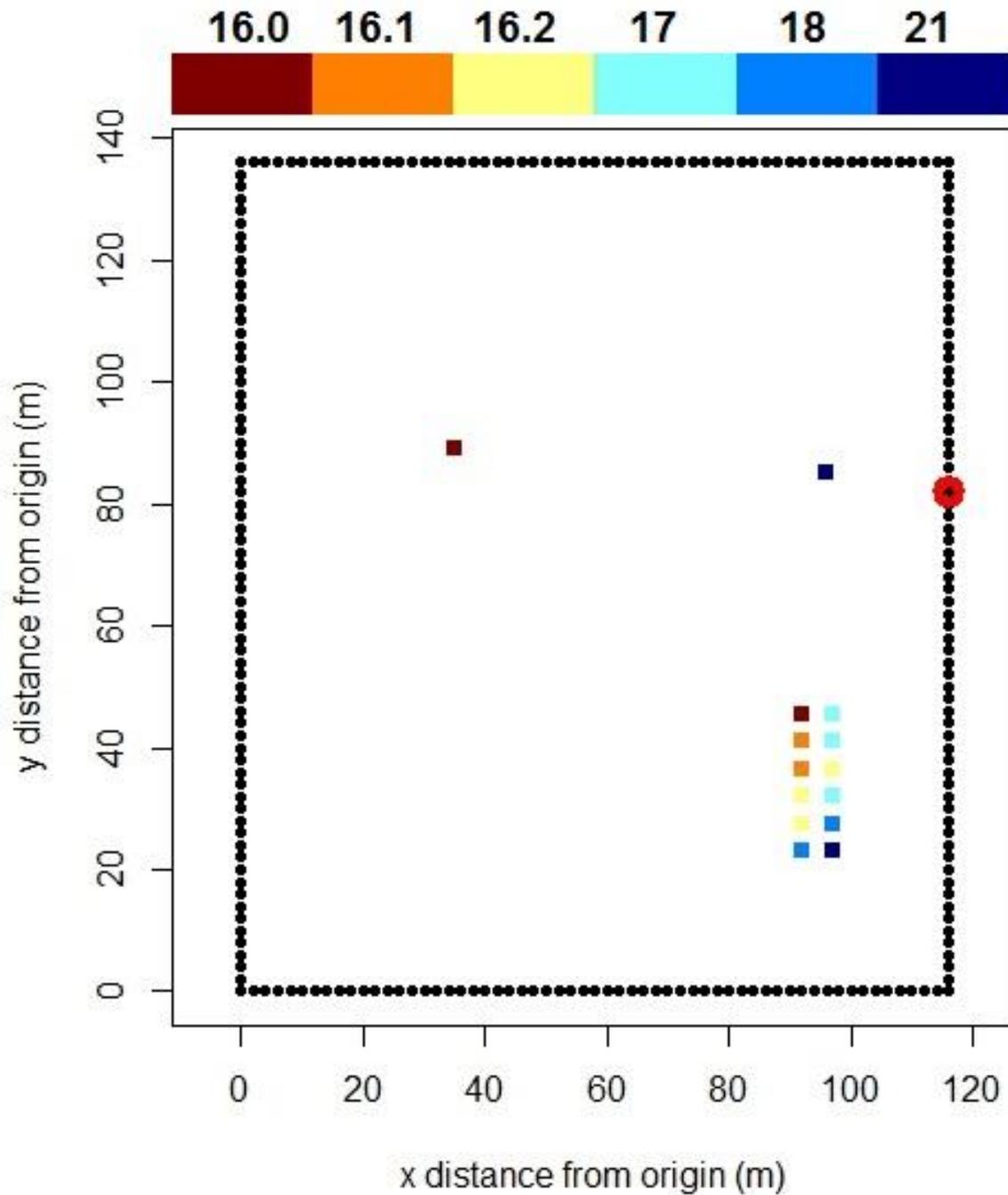
Summary

For the three well pads modeled by the Gaussian dispersion model, there are locations where point or open path sensors with a 0.5 ppm detection limit can detect 150 scfh emissions from most sources on average in less than 12 hours for most sources and within 48 hours for all sources based on local meteorological data. The 95th percentile time between detection, which accounts for periods of persisting unfavorable meteorological conditions, is only a few days for most sources with a maximum of 8 days for the worst source at the Bakken site. This suggests that the MDC leak detection systems will be able to quickly detect large leaks at most sites with only a single sensor. The detection time for 6 scfh leaks was poorer for 0.5 ppm detection limit sensors – in the Eagle Ford and Bakken sites, there were no evaluated receptors where all sources could be detected. In the Fayetteville, there was a receptor where a 0.5 ppm point sensor could detect most emissions within 3 days and the worst source within 18 days, which is adequate performance for detecting small leaks considering that they typically would only be identified within 6 months if sites were inspected semi-annually. Improving the detection limit to 0.1 ppm results in substantially improved performance for detecting 6 scfh leaks with most sources detected within a few days, but at the Bakken and Eagle Ford sites there are some sources that have long average detection time up to 35 days even at the best receptor locations. This is caused by some sources having either a long distance or large relative height difference from any receptor. At these sites, it may be necessary to employ multiple sensors to detect small emissions from all sources.

This dispersion model provides a first approximation of concentration enhancements based on Gaussian dispersion, but actual concentrations are harder to predict. First, the background concentration can vary radically depending on the stability class and local emission sources. In particular, there can be very high background methane concentrations during stable conditions in areas with many sources. For example, during the Barnett Shale campaign, background concentrations of 25 ppm compared to the global average of 1.8 ppm were observed in the nighttime when a descending atmospheric boundary layer compressed regional emissions from the previous day into a thin layer near the ground. Background emissions may also be affected by upwind offsite sources such as landfills or other O&G facilities. Therefore, it is critical that leak detection systems determine the concentration enhancement relative to a

dynamic background. This can be accomplished by evaluating the concentration time series, such as gradually increasing concentration during the nighttime that indicates boundary layer compression. Evaluating concentration in relationship to wind direction is also useful for determining if enhancement is caused by a discrete source rather than an increasing background, since the concentration should only vary by wind direction if it is from a source. The relationship of concentration and wind direction can also indicate if emissions are from onsite since the response of an offsite source to changing wind direction would be weaker than an onsite source.

One limitation of the model is that is relies on hourly meteorological data to calculate the detection time. Intra-hourly variation in wind direction would cause the model to be conservatively high in detection time since shifting winds would result in less time on average between conditions favorable for detection. Another limitation is the model does not account for either plume buoyancy or equipment downwash effects. If the buoyancy of a plume is known, then the model could be updated to account for buoyancy by modifying the vertical dispersion coefficient. This is feasible for vented sources from stacks since buoyancy could be estimated based on temperature and pressure, but impractical for equipment leaks since pressurized gas could escape with varying buoyancy depending on the leak point's orientation to the ground. The effect of equipment downwash would require more advanced dispersion modeling such as computational fluid dynamics. Typically, downwash would cause sensors to be less sensitive to relative height difference from a source. For example, emissions from tank hatches have been observed to turbulently mix down the sides of a tank, which leads to emissions being more quickly dispersed to ground-level.

Several improvements to the model are planned both prior to the MDC pilot deployments and after the leak detection systems start collecting data. First, the model will be updated to determine the combined performance of multiple networked sensors. For every modeled hour of meteorological data, detection will be assessed based on the measured enhancement at two or more receptors. If detecting small leaks with two 0.5 ppm detection limit sensors is substantially better than a single sensor, then multiple sensors may be more cost-effective than investing in detection limit improvements. Second, scenarios will be modeled with large offsite sources to assess how the relationship of wind direction and concentration can used to

distinguish offsite and onsite sources. Finally, measured concentration data from sensors at MDC pilot sites will be compared to modeled concentrations using higher frequency meteorological data. The comparison will provide information on how much dispersion deviates from Gaussian behavior and help evaluate observed concentrations in context of a site's potential emission sources. A critical step to reducing false alarms will be to determine the pattern of concentration enhancement from vented sources to distinguish normal emissions from leaks.

Continuous leak detection of methane emissions at O&G sites is an emerging technology with enormous potential to reduce emissions. Although the remaining challenges are great, this work has demonstrated that first generation leak detection systems will be able to quickly detect large leaks that are responsible for a large fraction of emissions. Modest improvements to detection limits will allow these systems to detect much smaller leaks and support the rapid mitigation of fugitive emissions. In the near future, cost-effective, continuous leak detection systems will be the foundation of leak detection and repair programs to reduce the climate impacts of the oil and gas industry.

CHAPTER 4 REFERENCES

1. Allen, D. T. *et al.* Measurements of methane emissions at natural gas production sites in the United States. *Proc. Natl. Acad. Sci.* **110,** 17768–17773 (2013).

2. Mitchell, A. L. *et al.* Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results. *Environ. Sci. Technol.* **49,** 3219–3227 (2015).

3. Subramanian, R. *et al.* Methane Emissions from Natural Gas Compressor Stations in the Transmission and Storage Sector: Measurements and Comparisons with the EPA Greenhouse Gas Reporting Program Protocol. *Environ. Sci. Technol.* **49,** 3252–3261 (2015).

4. Lyon, D. R. *et al.* Constructing a Spatially Resolved Methane Emission Inventory for the Barnett Shale Region. *Environ. Sci. Technol.* **49,** 8147–8157 (2015).

5. Zavala-Araiza, D. *et al.* Reconciling divergent estimates of oil and gas methane emissions. *Proc. Natl. Acad. Sci.* **112,** 15597–15602 (2015).

6. Eastern Research Group. *City of Fort Worth Natural Gas Air Quality Study.* (City of Fort Worth, 2011).

7. Allen, D. T. *et al.* Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Pneumatic Controllers. *Environ. Sci. Technol.* **49,** 633–640 (2015).

8. Lyon, D. R. *et al.* Aerial surveys of elevated hydrocarbon emissions from oil and gas production sites. *Environ. Sci. Technol.* (2016). doi:10.1021/acs.est.6b00705

9. *Proposed Rule: Oil and Natural Gas Sector: Emission Standards for New and Modified Sources.* **80 FR 56593,** (2015).

10. United States Environmental Protection Agency. *Alternative Work Practics to Detect Leaks from Equipment. 40 CFR Parts 60, 63, and 65*

11. Albertson, J. D. *et al.* A Mobile Sensing Approach for Regional Surveillance of Fugitive Methane Emissions in Oil and Gas Production. *Environ. Sci. Technol.* **50,** 2487–2497 (2016).

12. Leak Surveys, Inc. Available at: http://www.leaksurveysinc.com/.

13. Kairos Aerospace. Available at: http://kairosaerospace.com/leak-surveyor.

14. Physical Sciences, Inc. Available at: http://www.psicorp.com/.

15. Bridger Photonics. Available at: www.bridgerphotonics.com.

16. Turner, A. J. *et al.* Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data. *Atmospheric Chem. Phys. Discuss.* **15,** 4495–4536 (2015).

17. Kemp, C. E., Ravikumar, A. P. & Brandt, A. R. Comparing Natural Gas Leakage Detection Technologies Using an Open-Source 'Virtual Gas Field' Simulator. *Environ. Sci. Technol.* (2016). doi:10.1021/acs.est.5b06068

18. United States Department of Energy Advanced Research Projects Agency – Energy. *ARPA-E MONITOR Program* Available at: http://arpa-e.energy.gov/?q=programs/monitor.

19. Environmental Defense Fund. *Methane Detectors Challenge* Available at: https://www.edf.org/energy/natural-gas-policy/methane-detectors-challenge.

20. R Development Core Team. *R: A language and environment for statistical computing.* (R Foundation for Statistical Computing, 2008).

21. Allen, D. T. & Durrenberger, C. J. Gaussian Plume Modeling. (2015). Available at: http://dept.ceer.utexas.edu/ceer/che357/PDF/Lectures/gaussian_plume_modeling.pdf.

22. Pasquill, F. The estimation of the dispersion of windborne material. *Meteorol Mag* **90,** 33–49 (1961).

23. Turner, D. B. *Workbook of atmospheric dispersion estimates.* (United States Environmental Protection Agency, 1969).

24. Arnott, P. Solar calculator for solar zenith angle. (2016). Available at: www.patarnott.com/atms360/general/SimpleSolarAngle.xls.

25. Iowa Environmental Mesonet. Automated Surface Observing System Network. (2016). Available at: https://mesonet.agron.iastate.edu/request/download.phtml.

CHAPTER 4 APPENDIX

Figure S1. Eagle Ford Shale meteorological data: Sensitivity of 95th percentile hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and distance of receptor from source
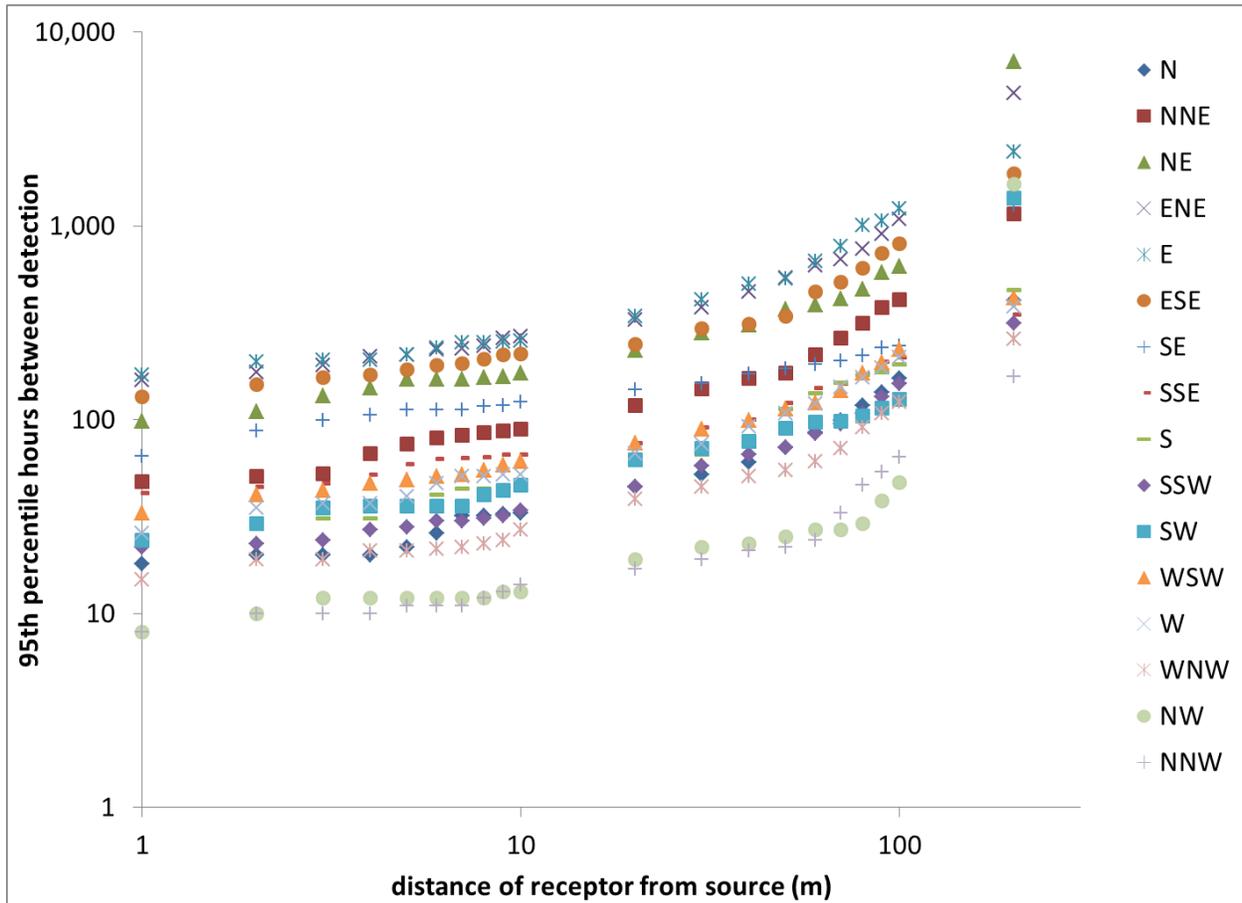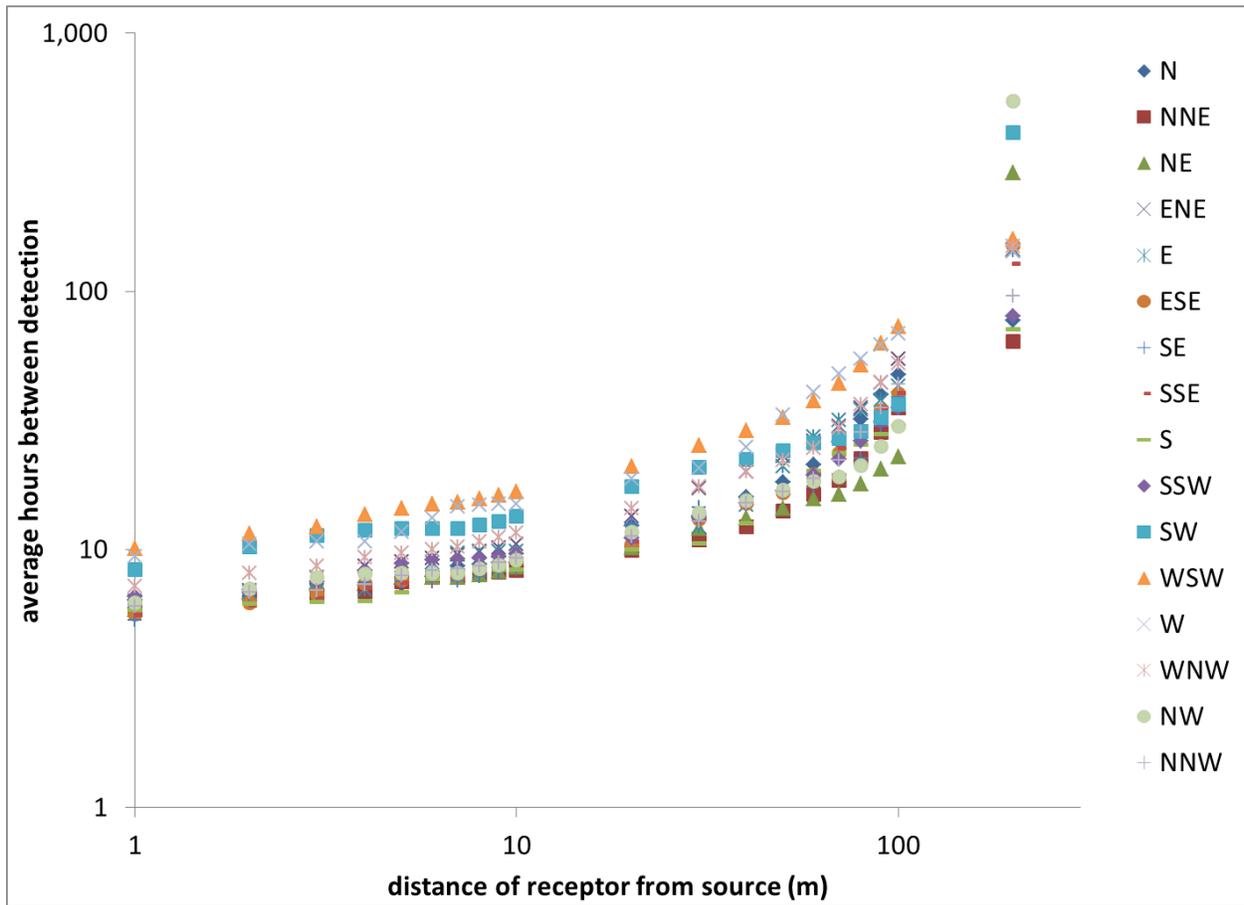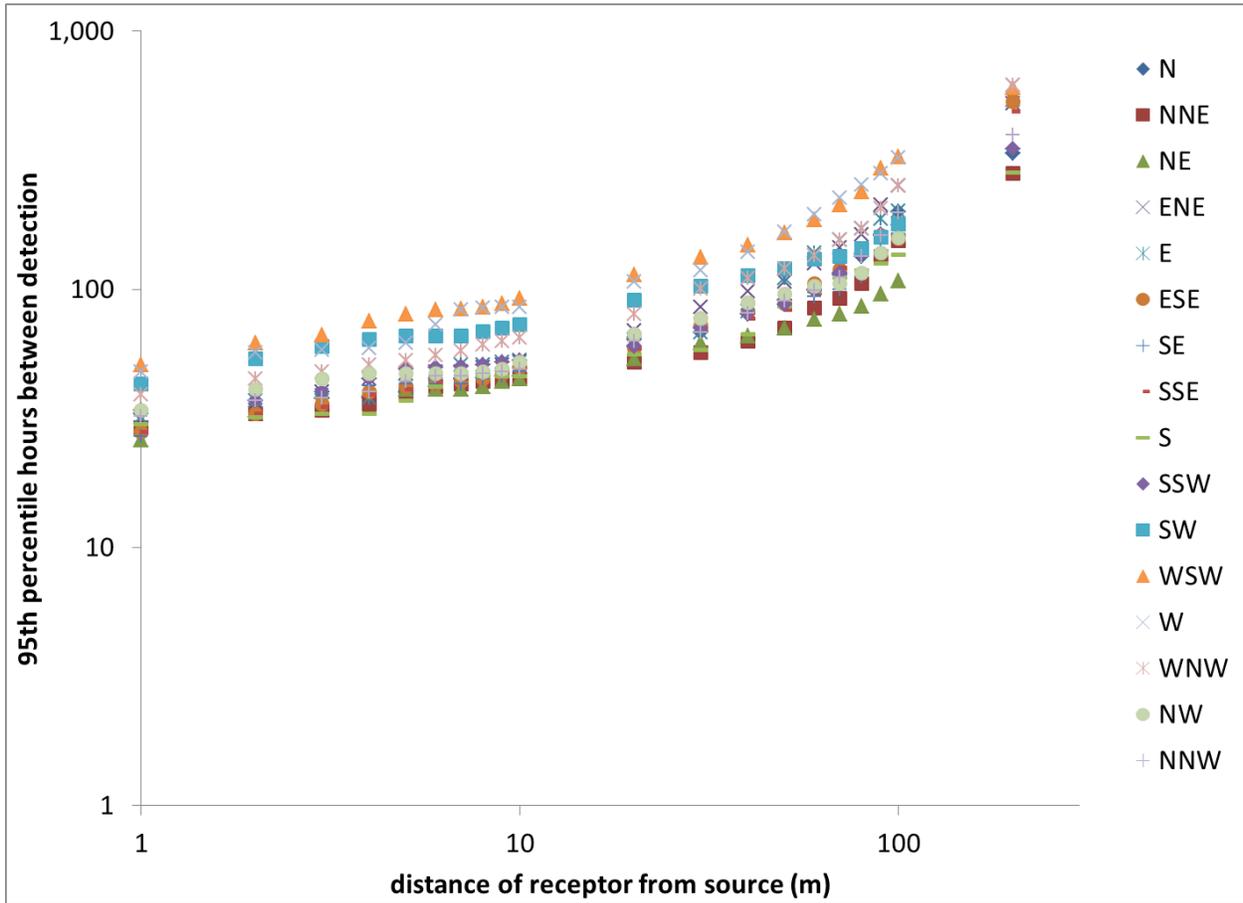
Figure S2. Eagle Ford Shale meteorological data: Sensitivity of average hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and relative height distance between a source and a receptor at 100 m distance
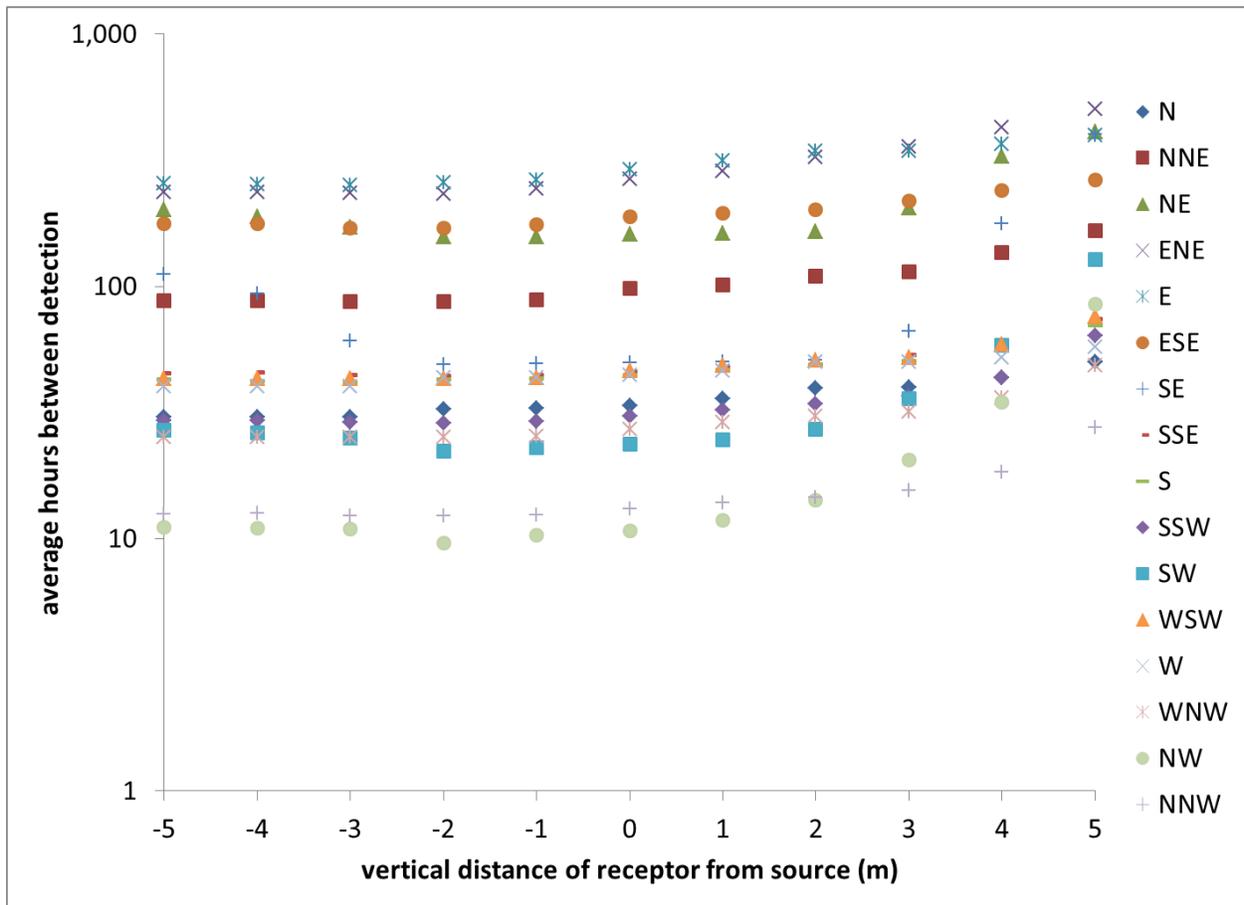
Figure S3. Fayetteville Shale meteorological data: Wind rose of Little Rock Air Force Base (LRF) meteorological data
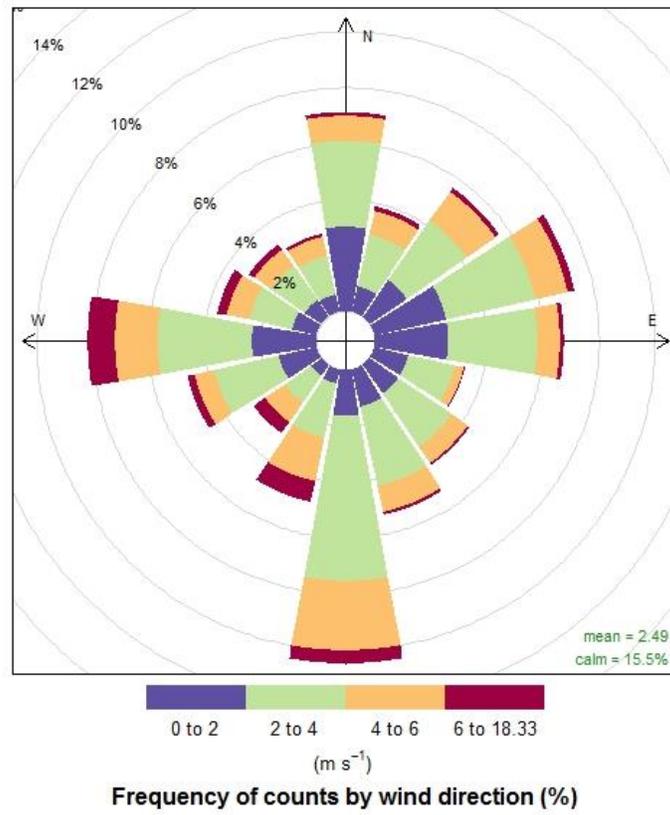
Figure S4. Bakken Shale meteorological data: Wind rose of Sloulin Field International Airport (ISN) meteorological data
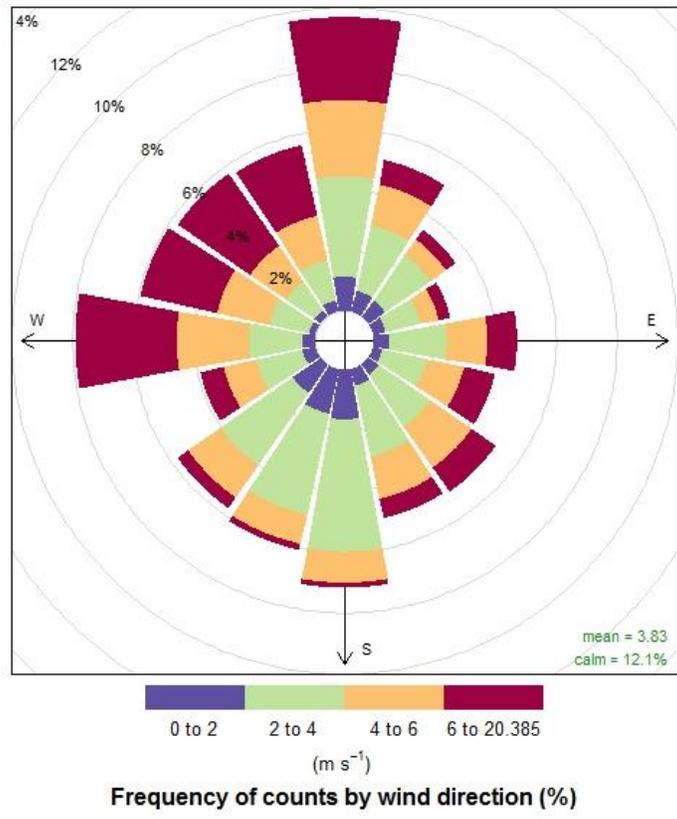
Figure S5. Fayetteville Shale meteorological data: Sensitivity of average hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and distance of receptor from source
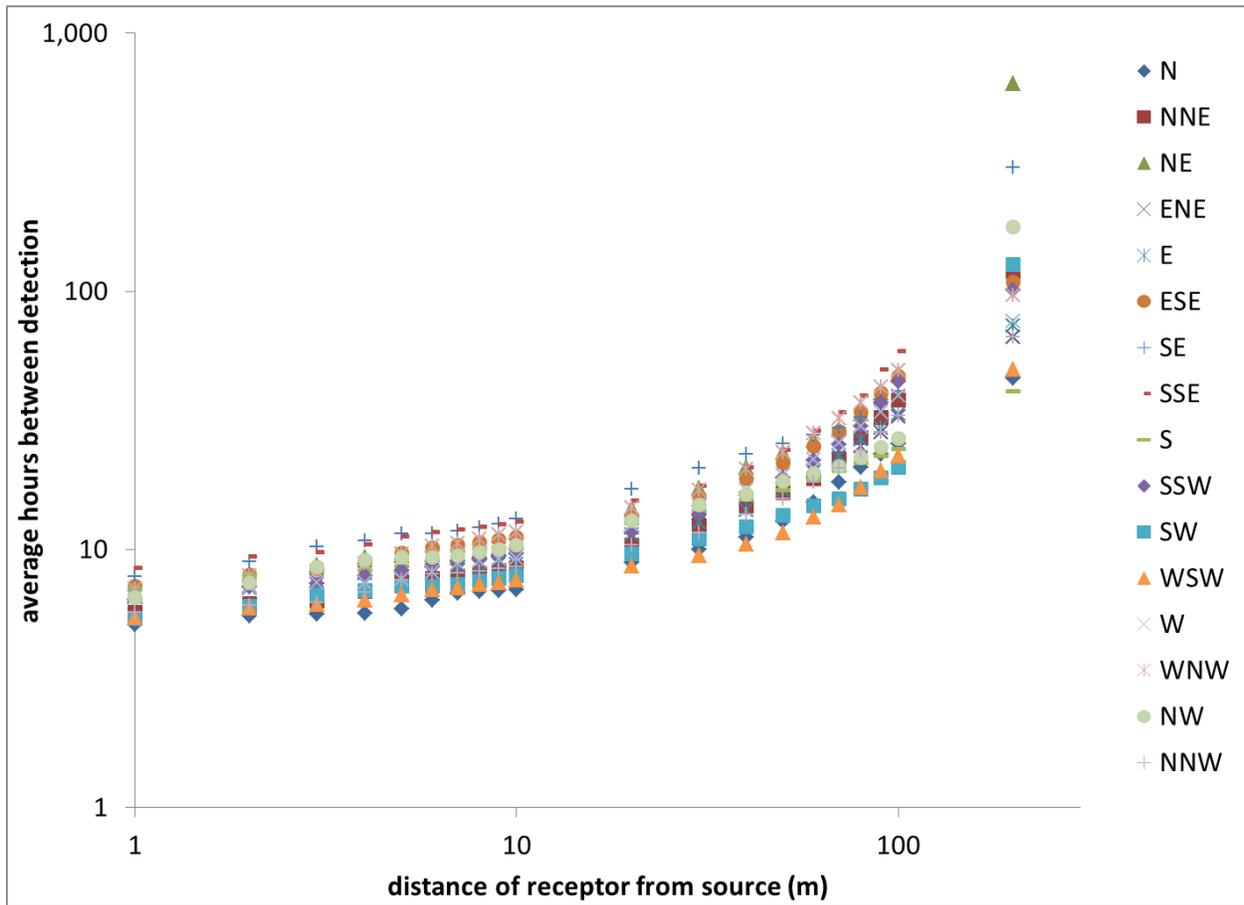
Figure S6. Fayetteville Shale meteorological data: Sensitivity of 95th percentile hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and distance of receptor from source
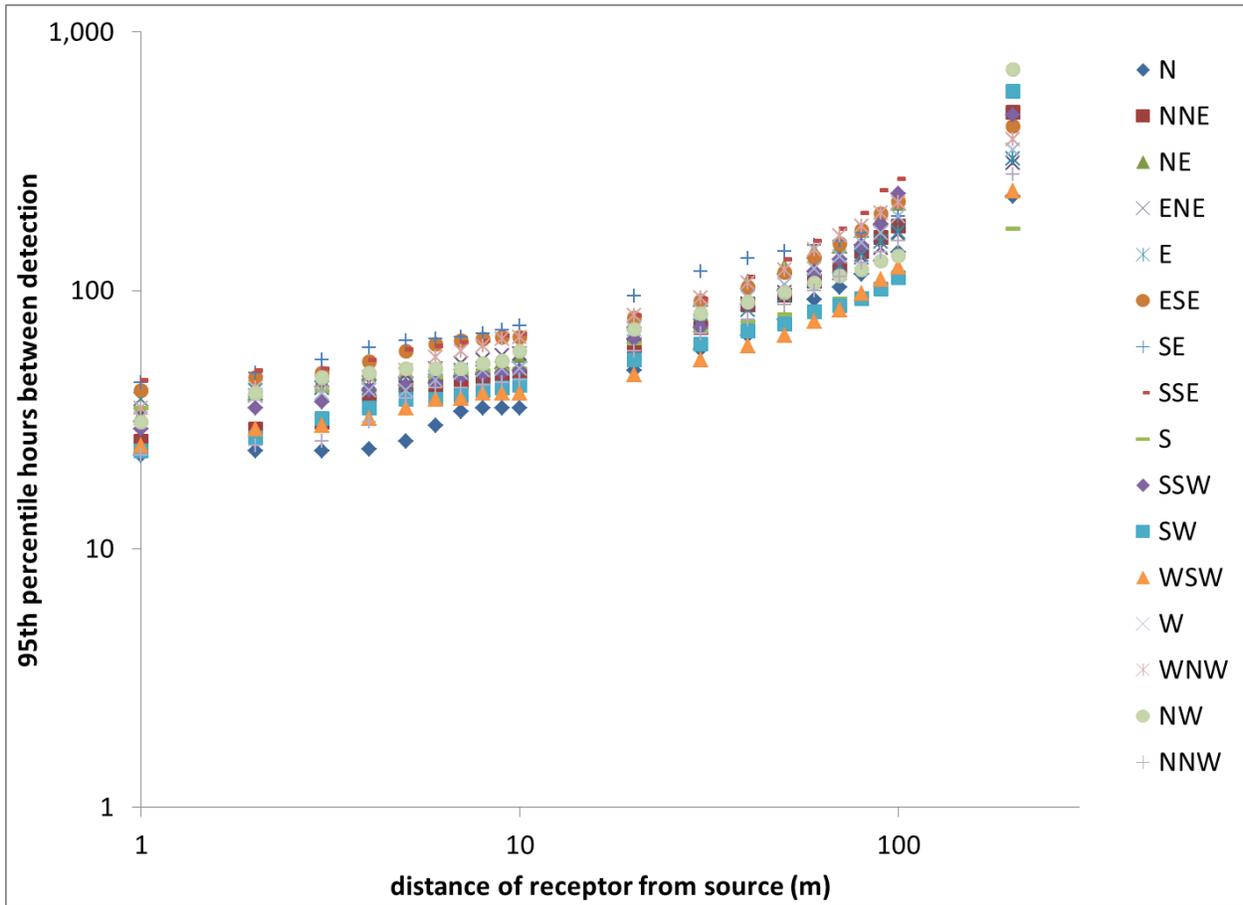
Figure S7. Fayetteville Shale meteorological data: Sensitivity of average hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and relative height distance between a source and a receptor at 10 m distance
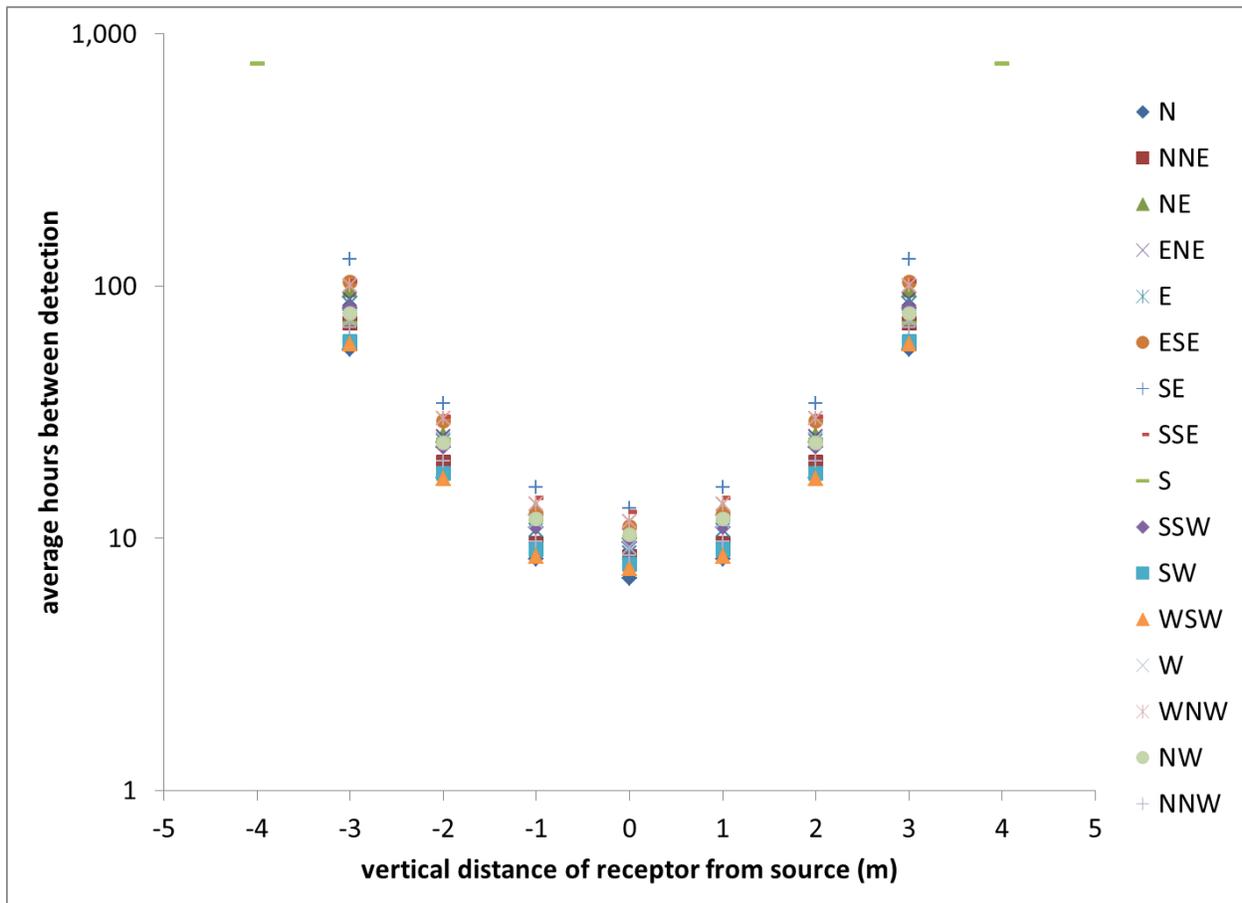
Figure S8. Fayetteville Shale meteorological data: Sensitivity of average hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and relative height distance between a source and a receptor at 100 m distance
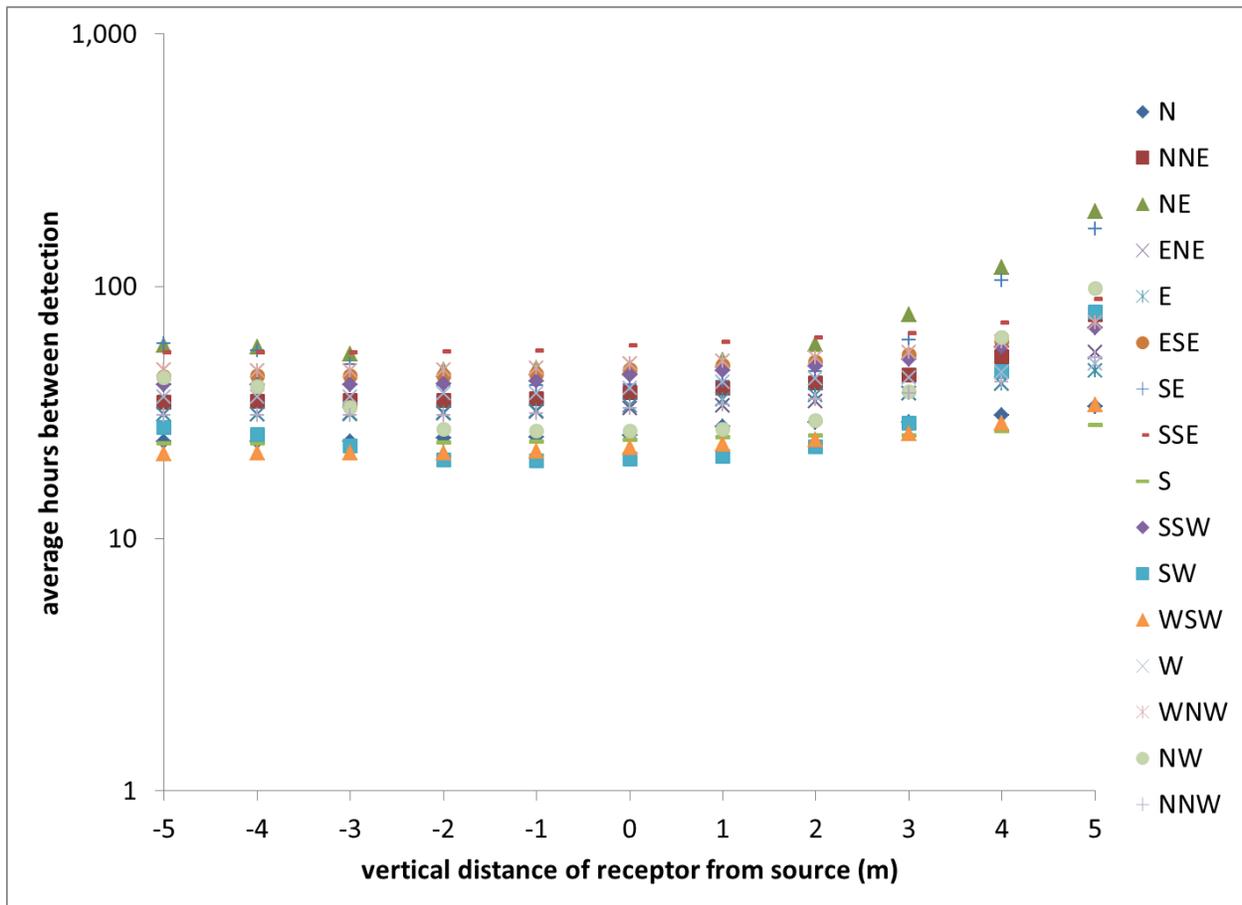
Figure S9. Bakken Shale meteorological data: Sensitivity of average hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and distance of receptor from source

Figure S10. Bakken Shale meteorological data: Sensitivity of 95th percentile hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and distance of receptor from source

Figure S11. Bakken Shale meteorological data: Sensitivity of average hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and relative height distance between a source and a receptor at 10 m distance
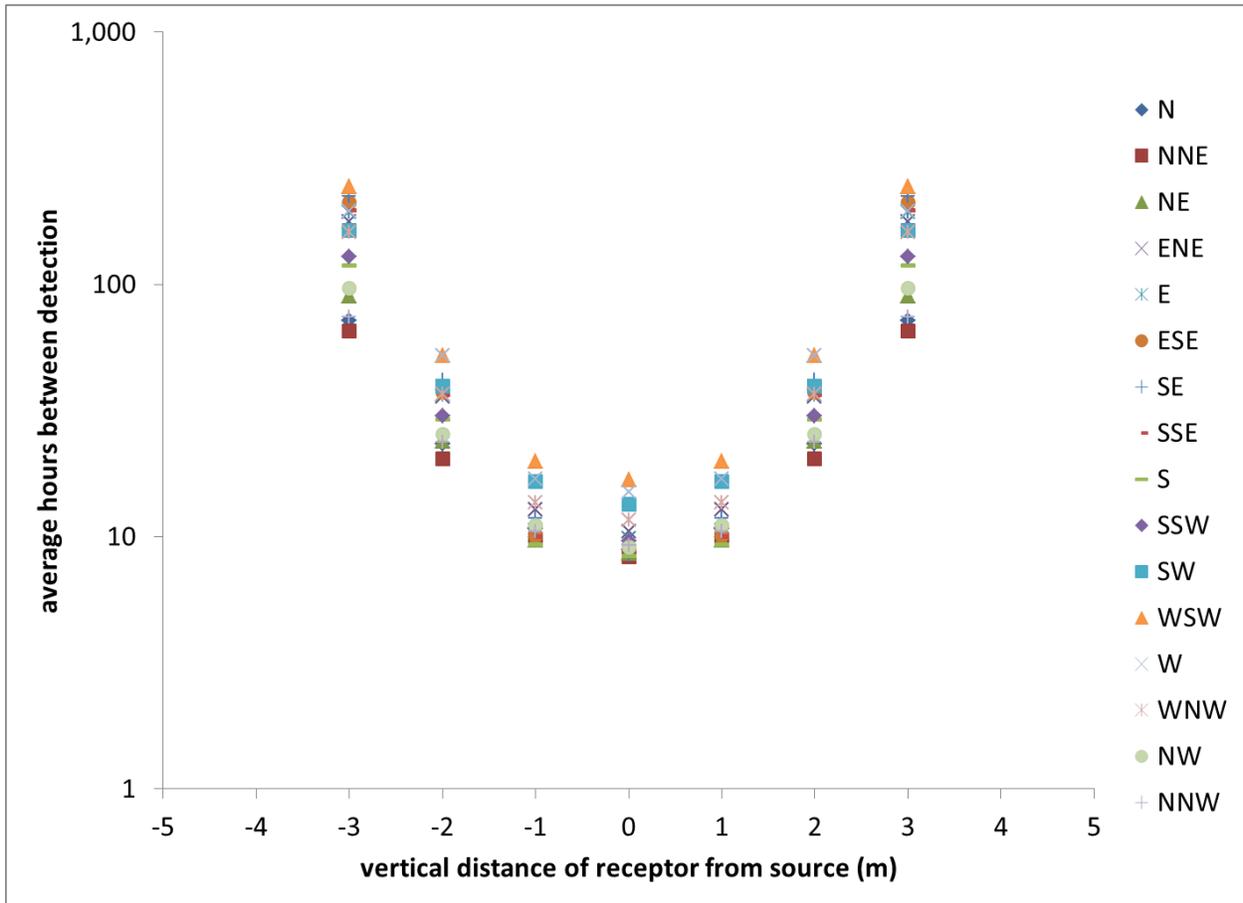
Figure S12. Bakken Shale meteorological data: Sensitivity of average hours between detection of a 150 scfh leak with a 0.5 ppm detection limit based on orientation and relative height distance between a source and a receptor at 100 m distance
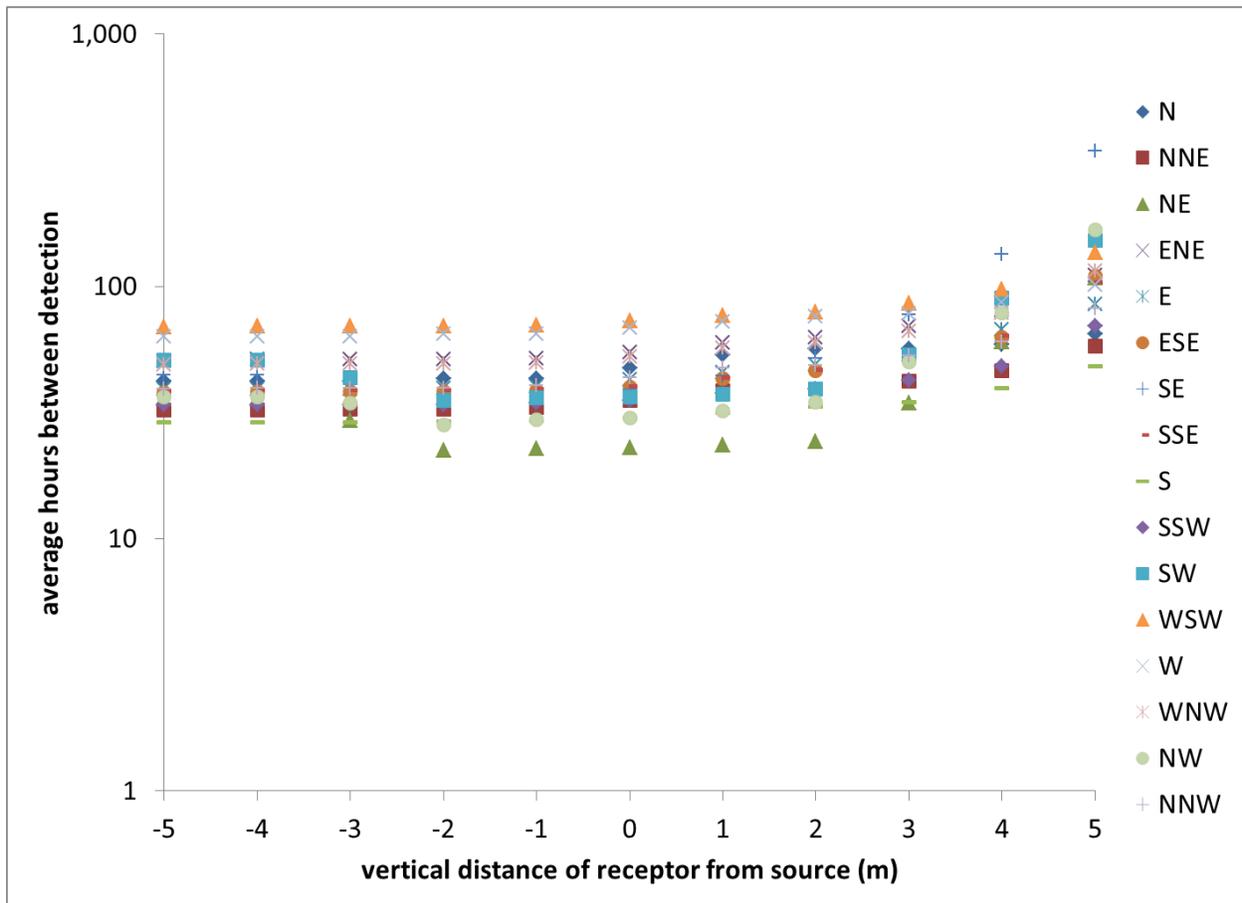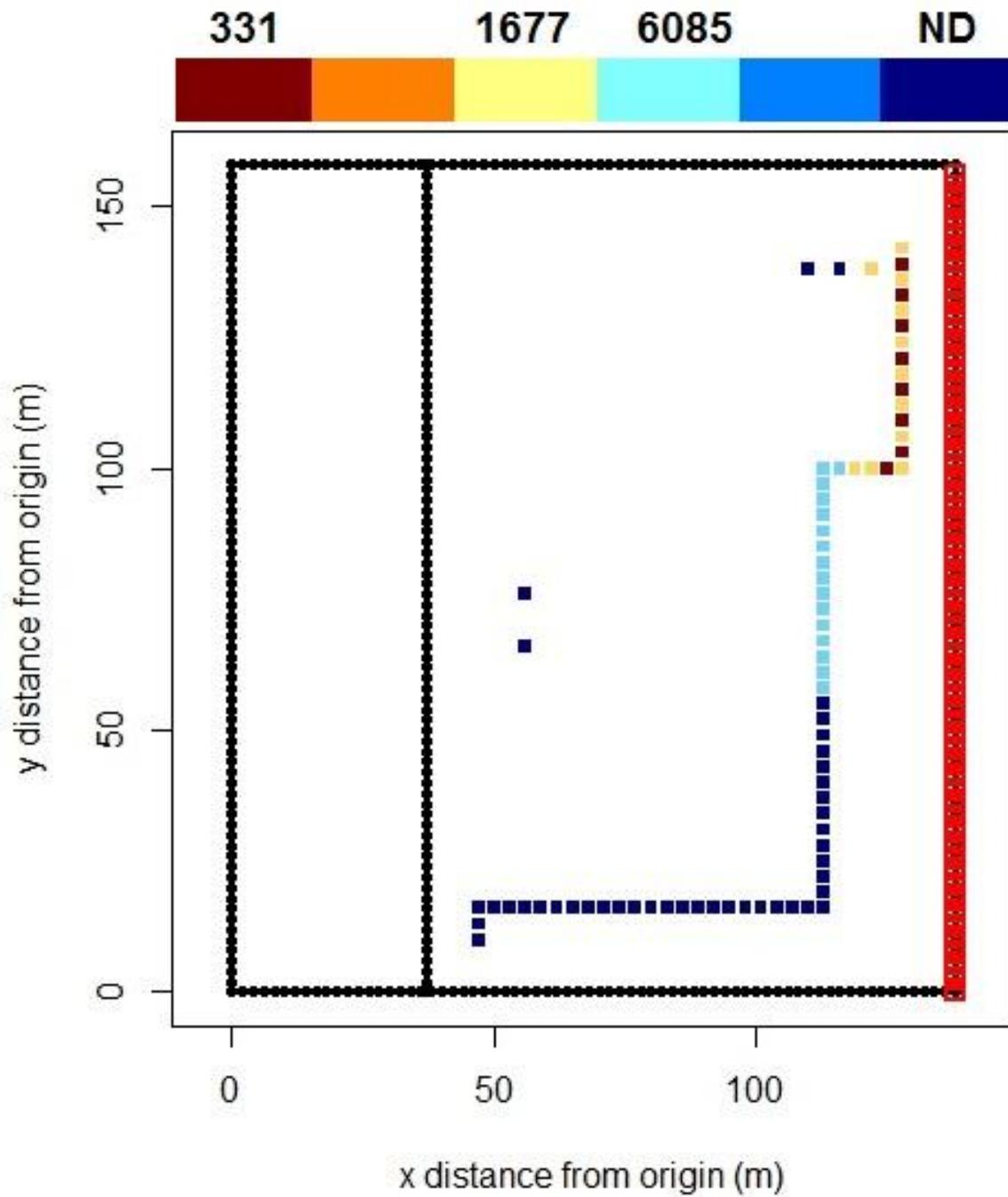
Figure S13. Eagle Ford well pad: Cumulative distribution function of hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit point sensor for every source-receptor combination (blue = average, green = 95th percentile) and the best receptor R366 (orange = average, red = 95th percentile)
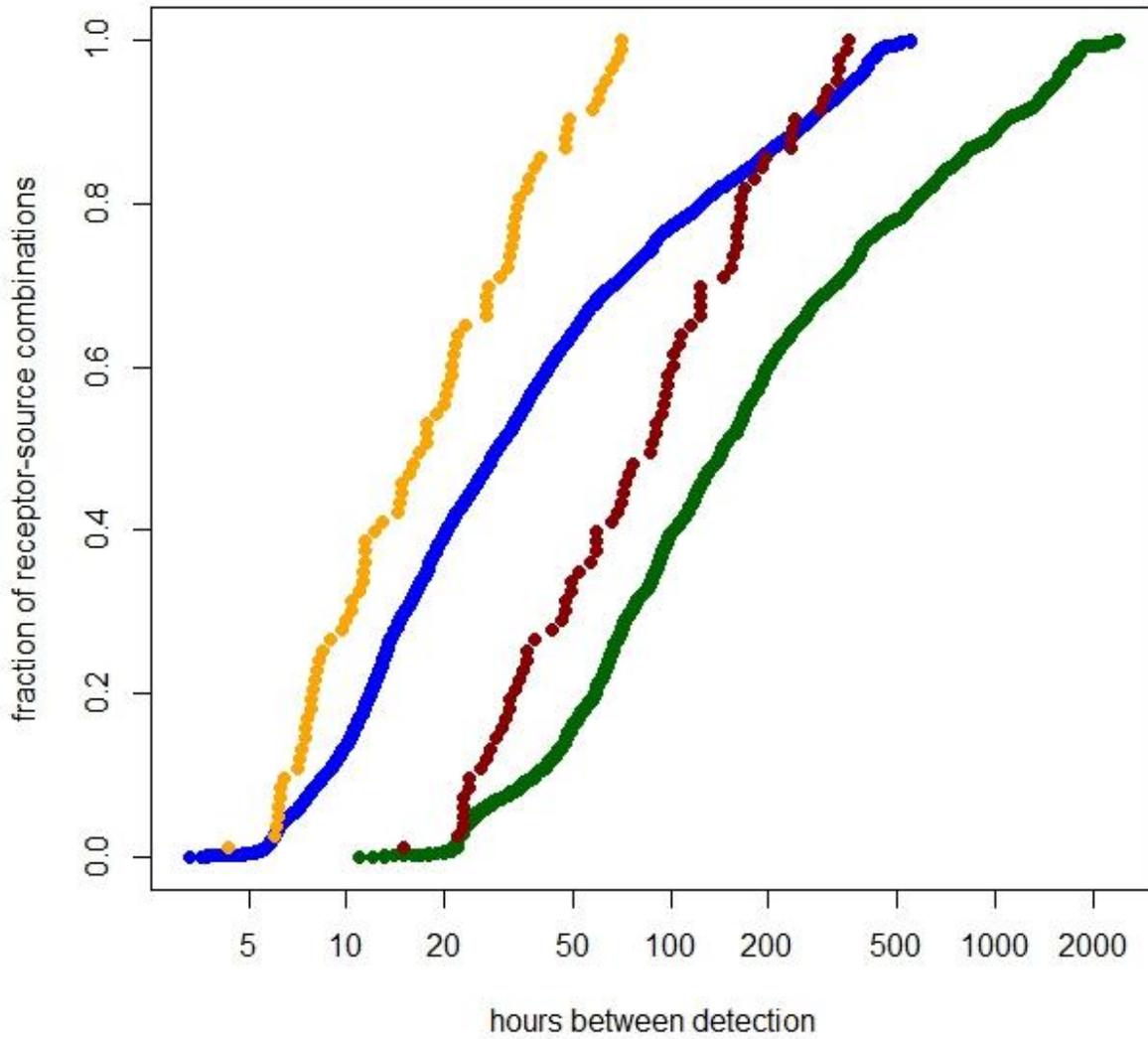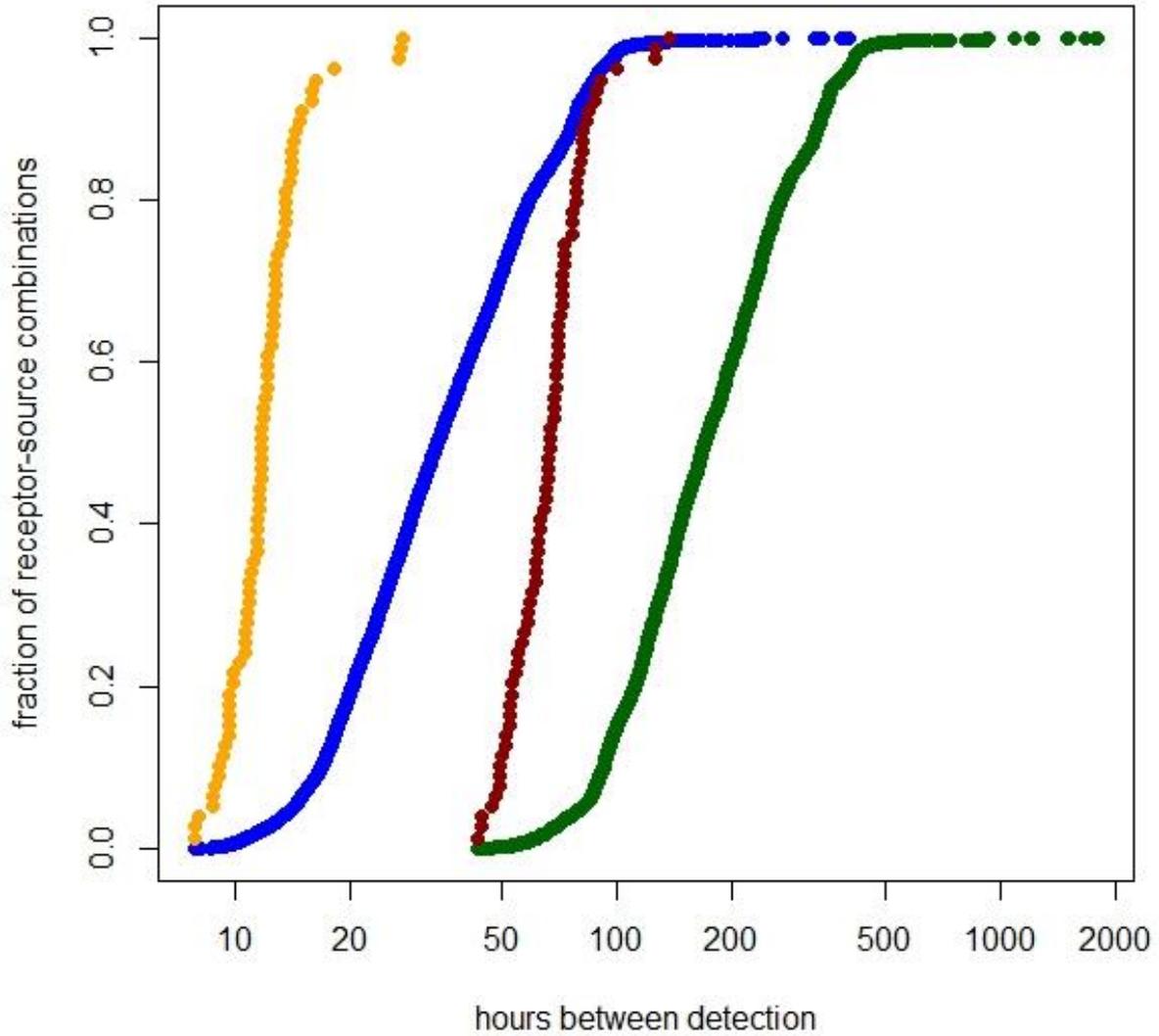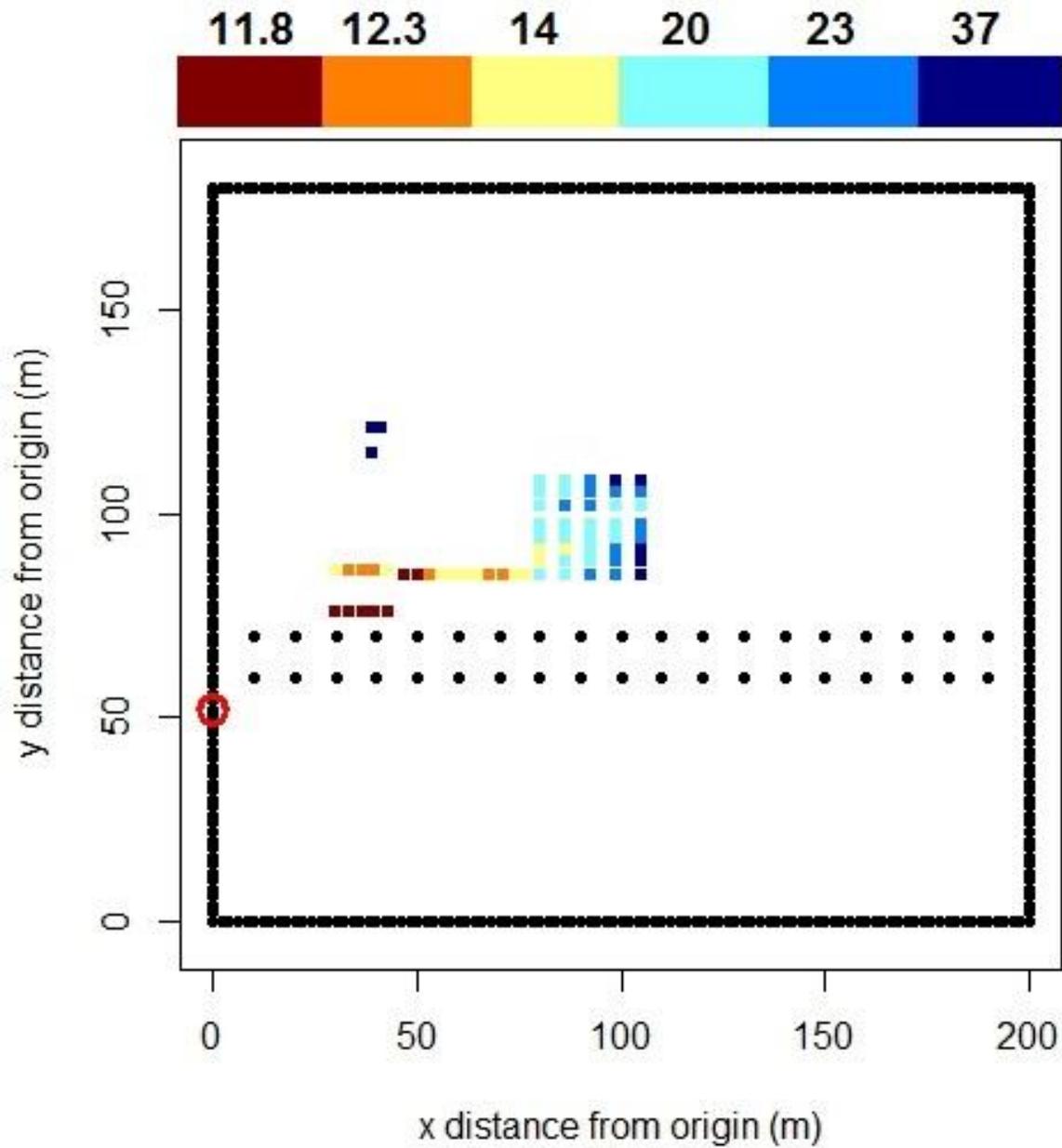
Figure S14. Eagle Ford well pad: Cumulative distribution function of hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit open path sensor for every source-path combination (blue = average, green = 95th percentile) and the best path E1 (orange = average, red = 95th percentile)
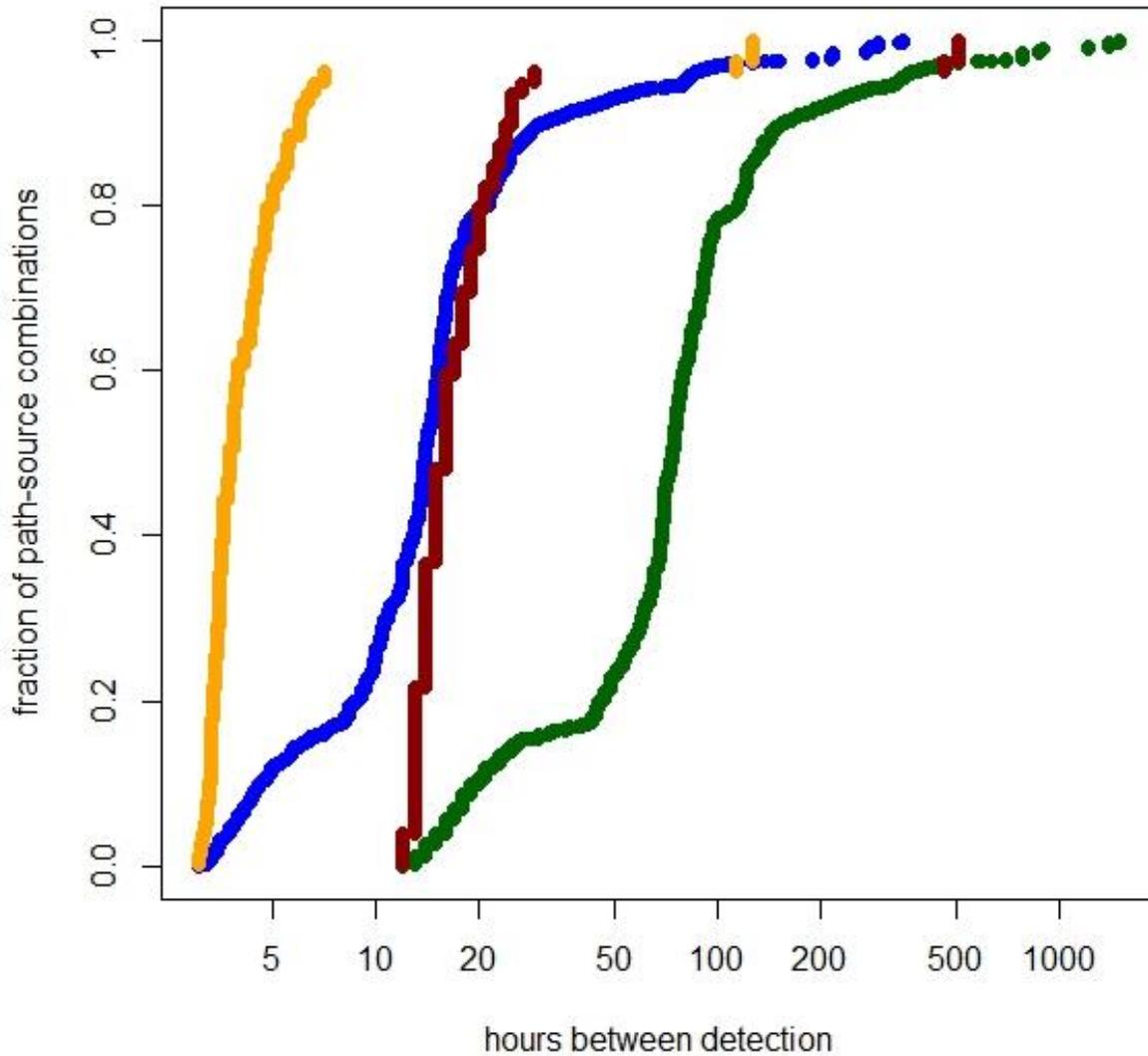
Figure S15. Eagle Ford well pad: Site-level average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.5 ppm detection limit point sensor for every receptor (1st, 2.5th, 5th, 10th, 25th, 50th, & 75th percentiles). Site-level metrics are based on the median of individual sources' average hours.
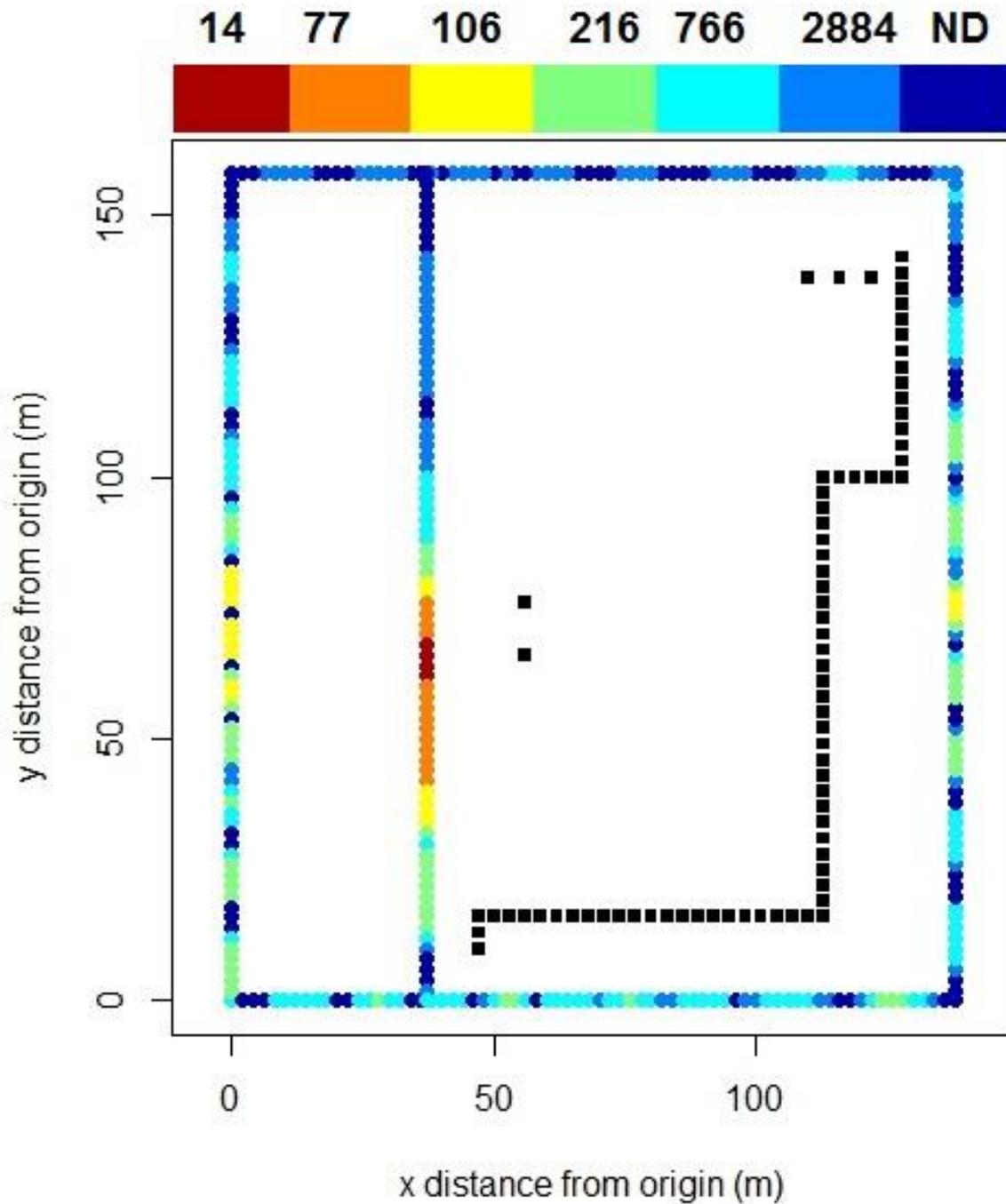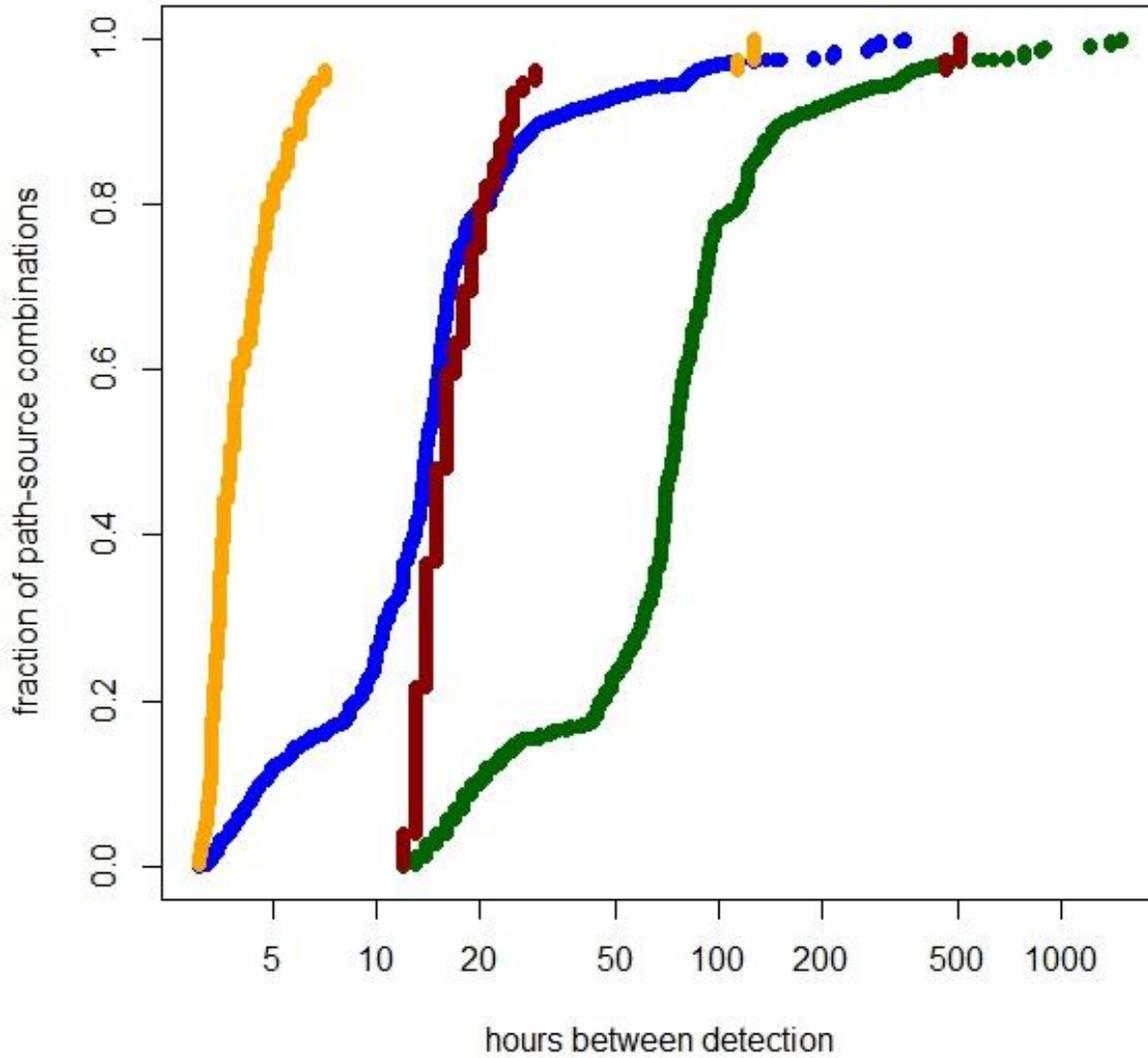
Figure S16. Eagle Ford well pad: Site-level average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.1 ppm detection limit point sensor for every receptor (1st, 2.5th, 5th, 10th, 25th, 50th, & 75th percentiles). Site-level metrics are based on the median of individual sources' average hours.
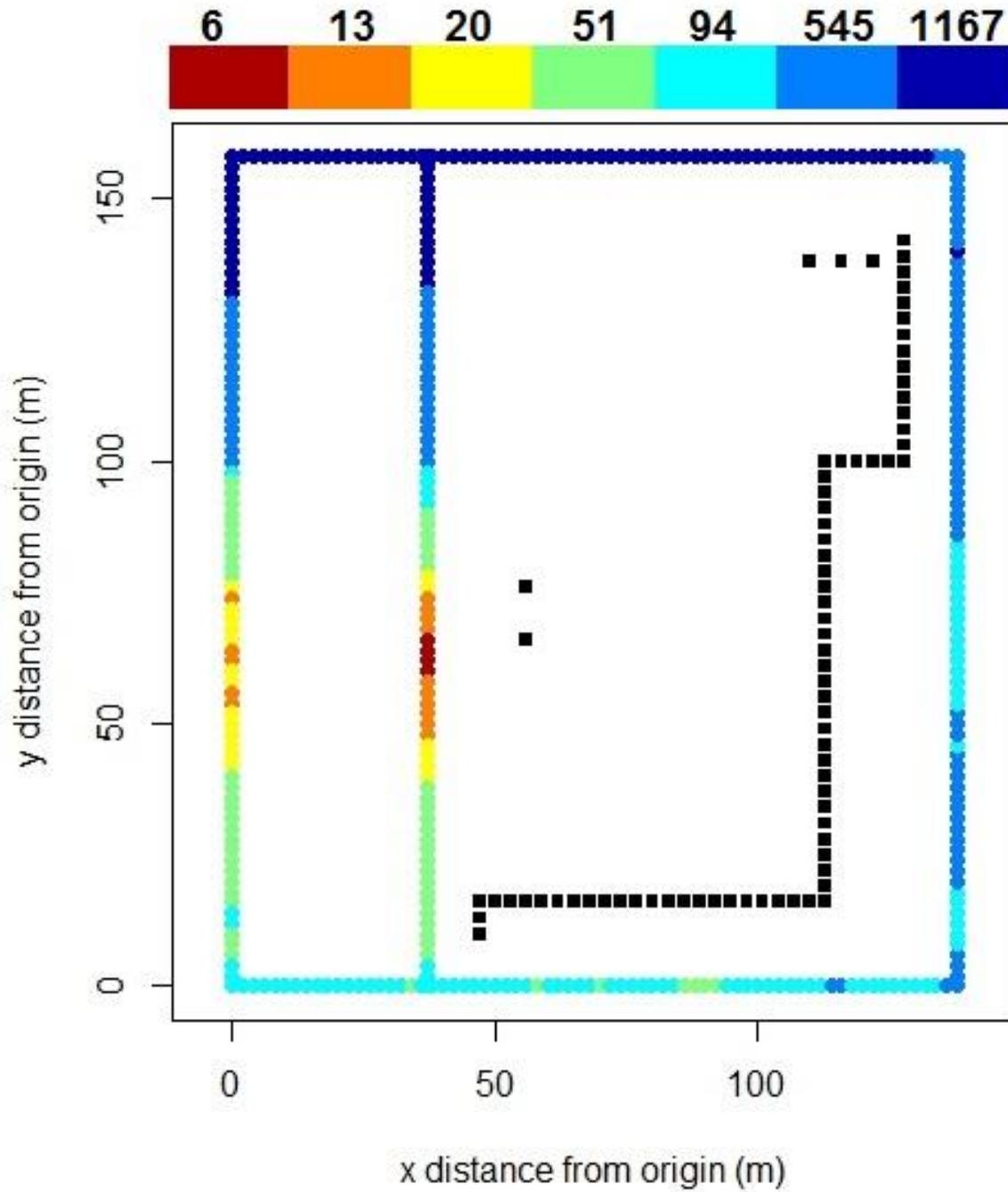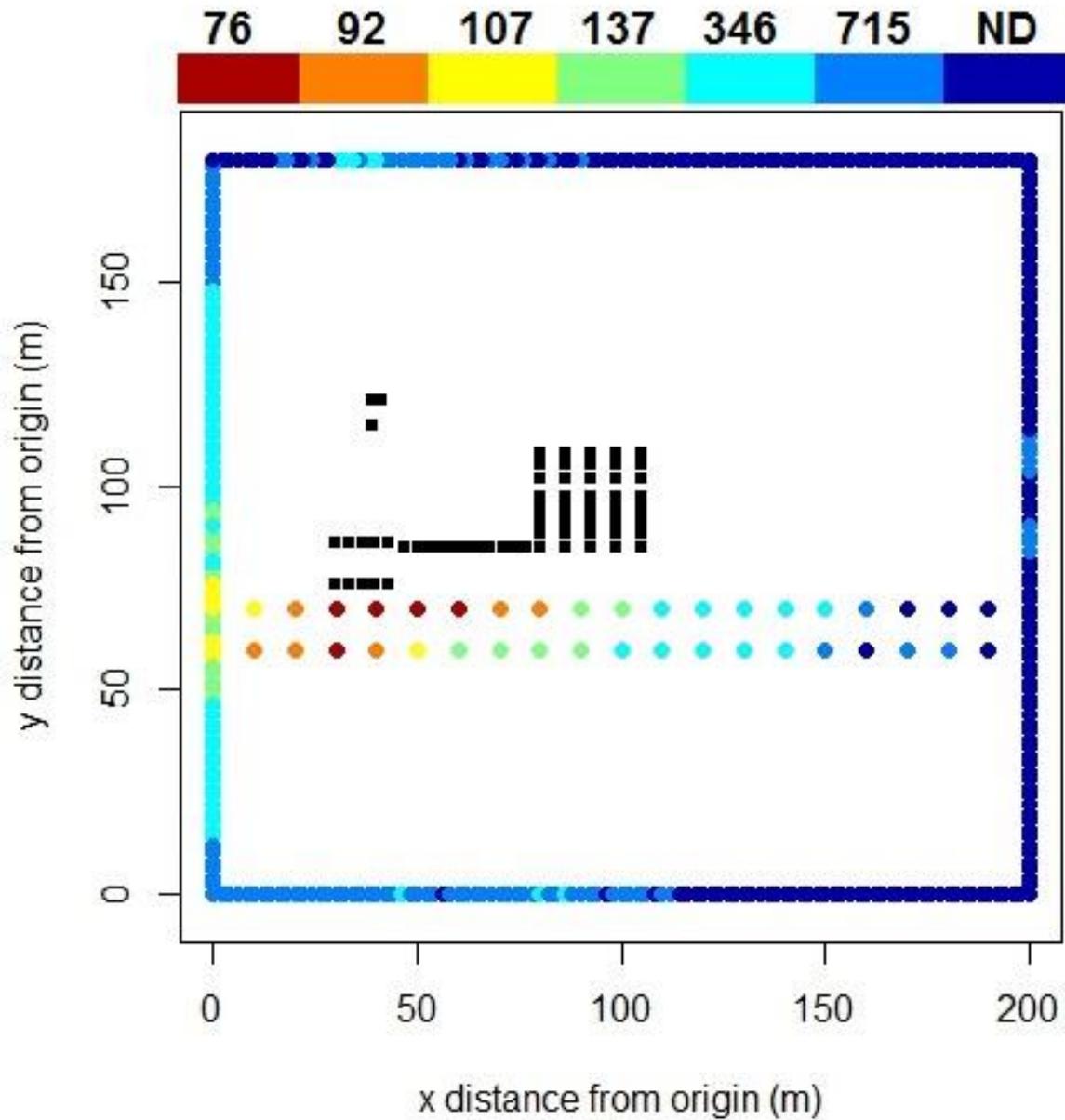
Figure S17. Eagle Ford well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources (10th, 25th, 50th, 75th, & 90th percentiles)with a 0.5 ppm detection limit point sensor at the best receptor (R82, highlighted in red).
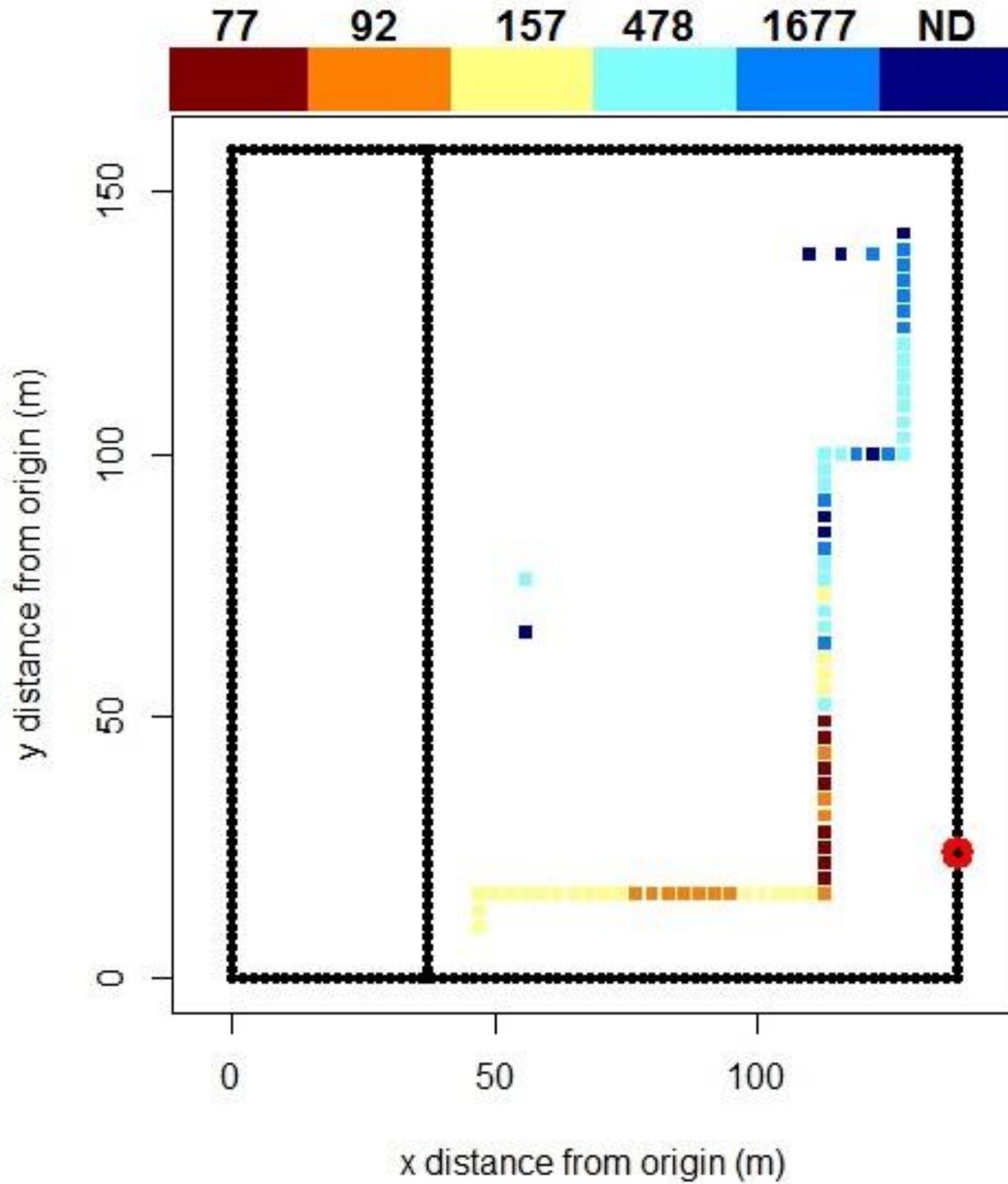
Figure S18. Eagle Ford well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources (10th, 25th, 50th, 75th, & 90th percentiles)with a 0.5 ppm detection limit open path sensor deployed across path E1 (highlighted in red).

Figure S19. Fayetteville well pad: Cumulative distribution function of hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit point sensor for every source-receptor combination (blue = average, green = 95th percentile) and the best receptor R366 (orange = average, red = 95th percentile)

Figure S20. Fayetteville well pad: Average hours between detection of a 150 scfh emission rate from individual sources (10th, 25th, 50th, 75th, & 90th percentiles) with a 0.5 ppm detection limit point sensor located at the best receptor located at the fence line (R355, highlighted in red)

Figure S21. Fayetteville well pad: Cumulative distribution function of hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit open path sensor for every source-path combination (blue = average, green = 95th percentile) and the best path I2 (orange = average, red = 95th percentile)
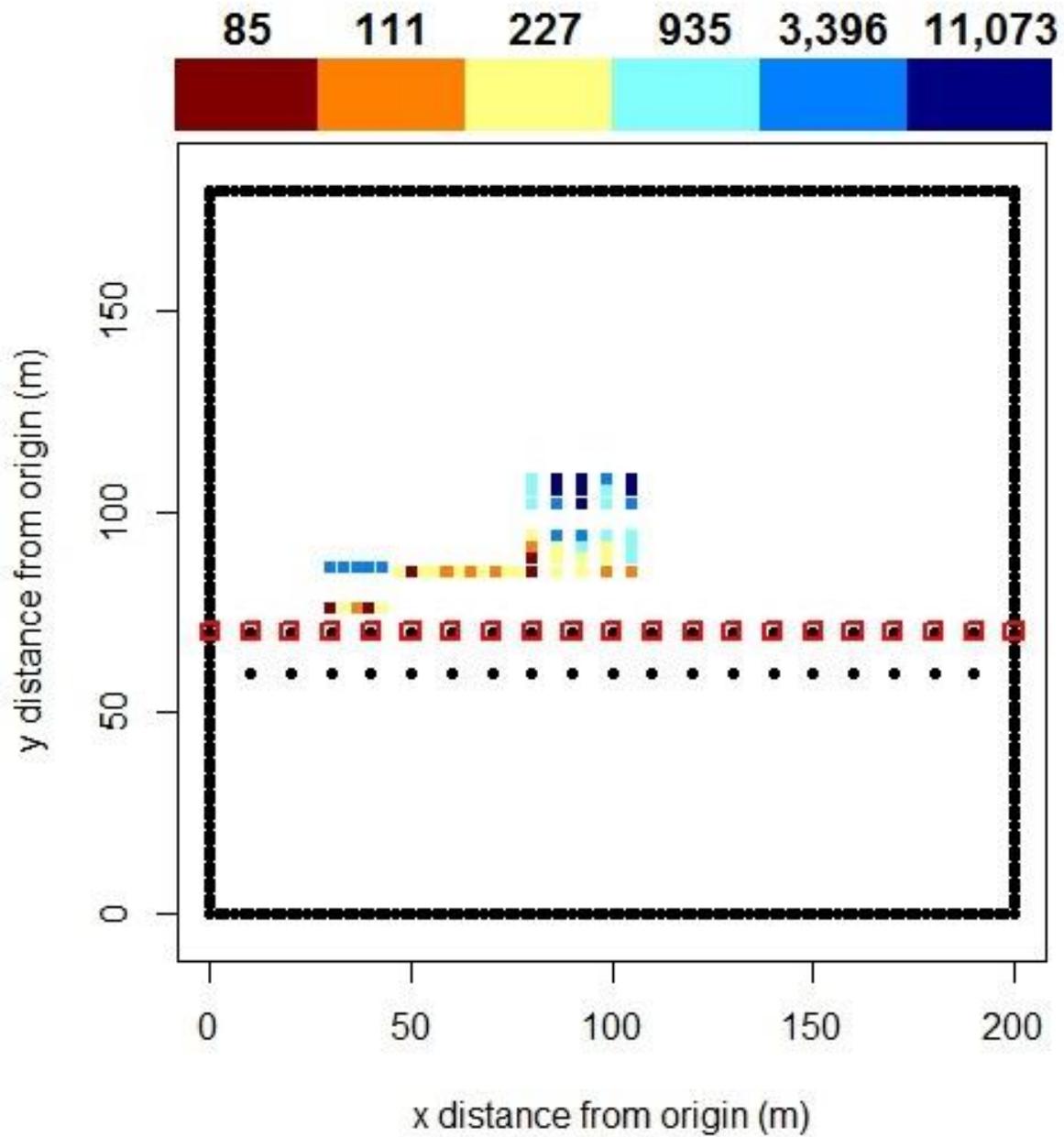
Figure S22. Fayetteville well pad: Site-level average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.5 ppm detection limit point sensor for every receptor (1st, 2.5th, 5th, 10th, 25th, 50th, & 75th percentiles). Site-level metrics are based on the median of individual sources' average hours.
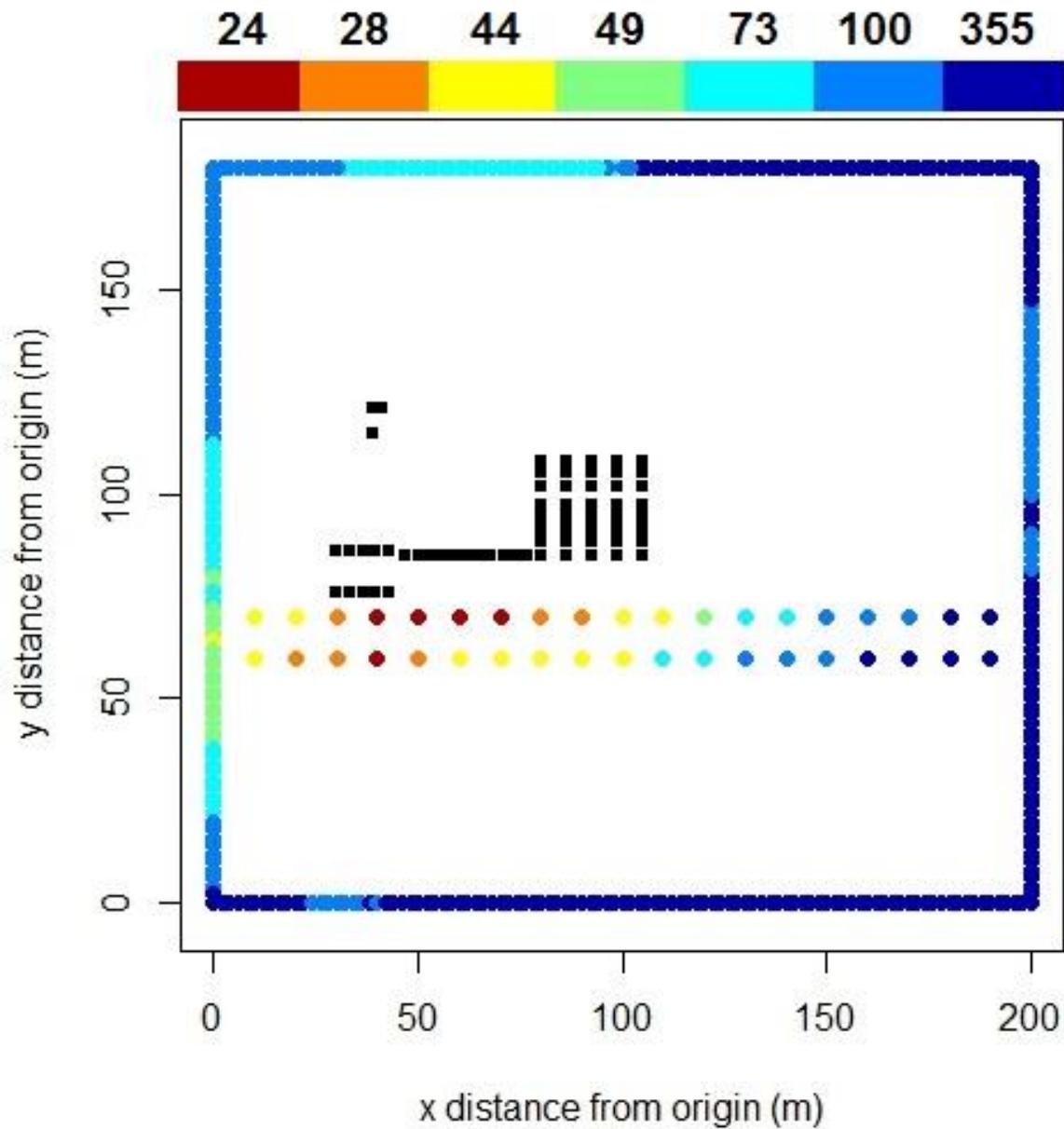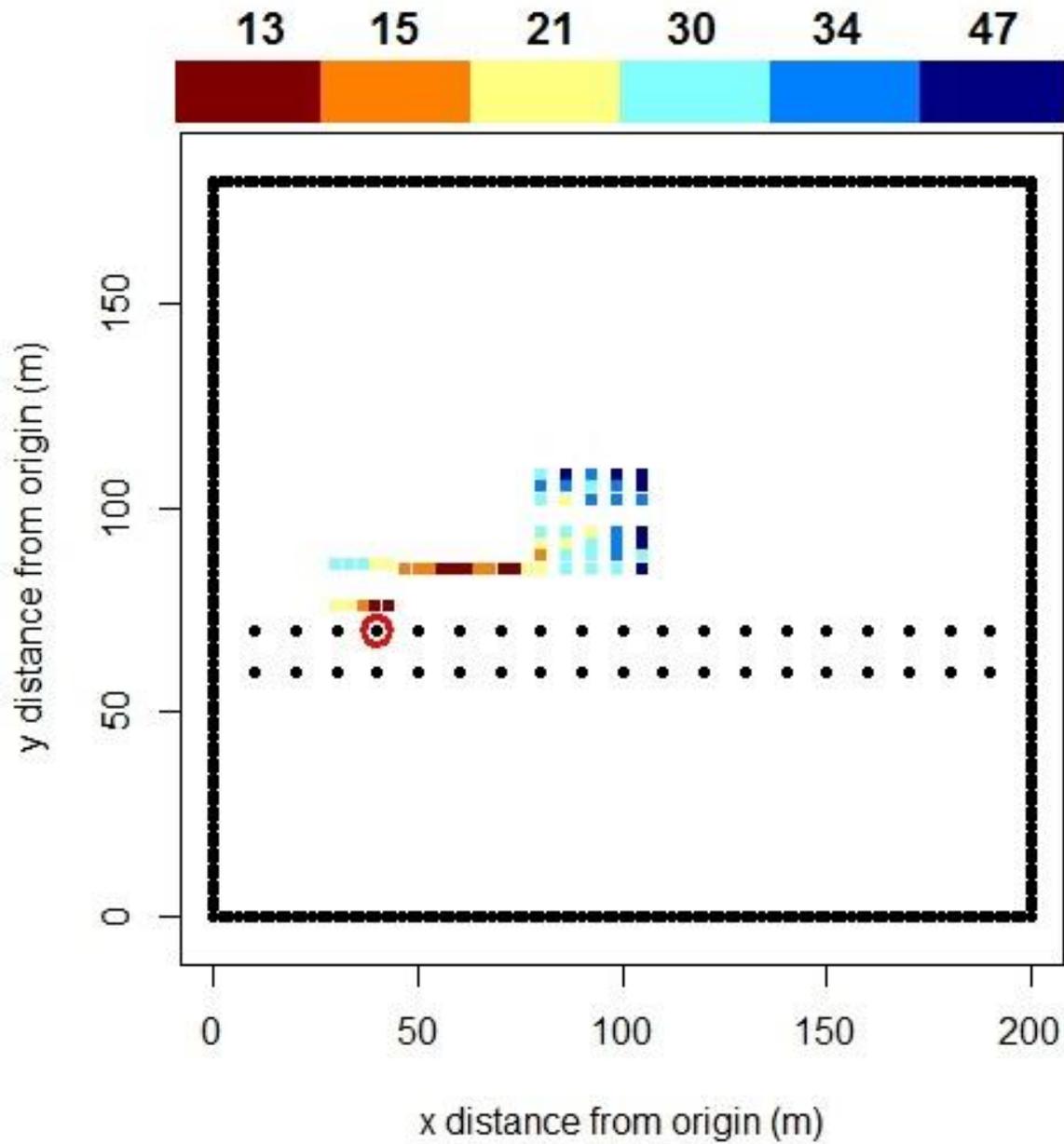
Figure S23. Fayetteville well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources (10th, 25th, 50th, 75th, & 90th percentiles)with a 0.5 ppm detection limit point sensor located at the best receptor located on the fence line (R344, highlighted in red).
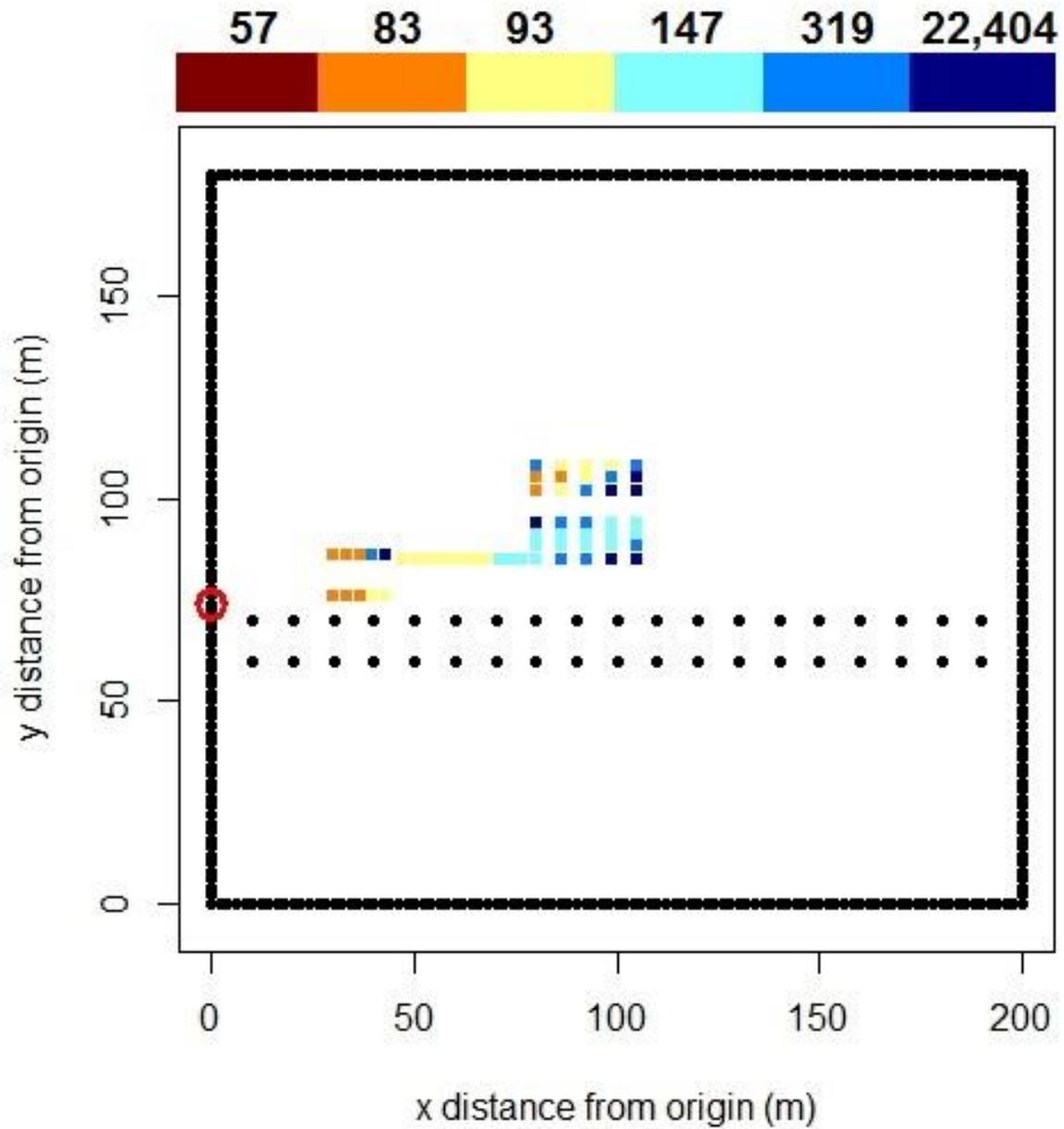
Figure S24. Fayetteville well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources (10th, 25th, 50th, 75th, & 90th percentiles)with a 0.5 ppm detection limit open path sensor deployed across the best path (l1, highlighted in red).

Figure S25. Fayetteville well pad: Site-level average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.1 ppm detection limit point sensor for every receptor (1st, 2.5th, 5th, 10th, 25th, 50th, & 75th percentiles). Site-level metrics are based on the median of individual sources' average hours.
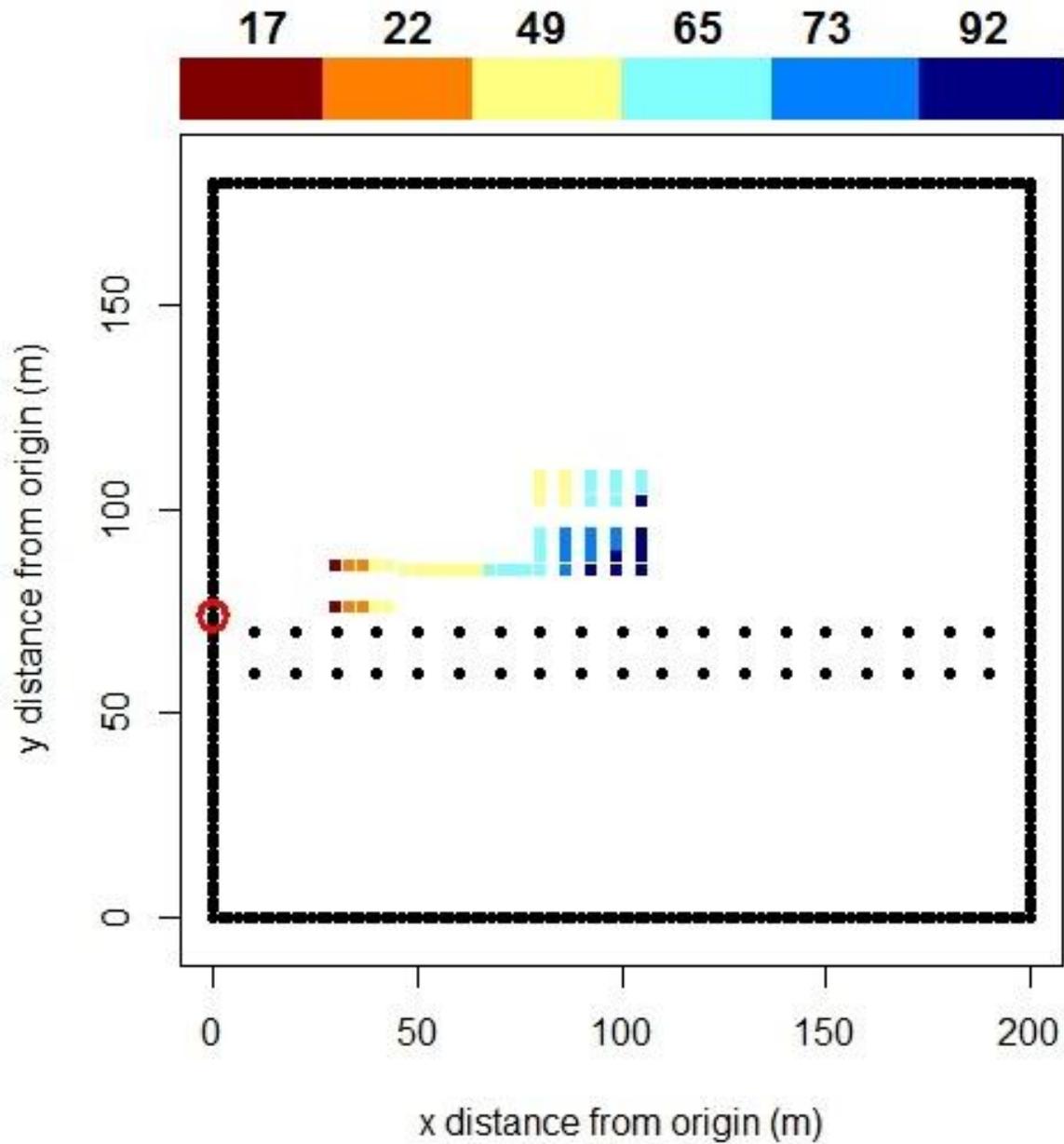
Figure S26. Fayetteville well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.1 ppm detection limit point sensor located at the best receptor (R433, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).

Figure S27. Fayetteville well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources (10th, 25th, 50th, 75th, & 90th percentiles)with a 0.1 ppm detection limit point sensor located at the best receptor located on the fence line (R344, highlighted in red).

Figure S28. Fayetteville well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.1 ppm detection limit open path sensor deployed across the best path (l1, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
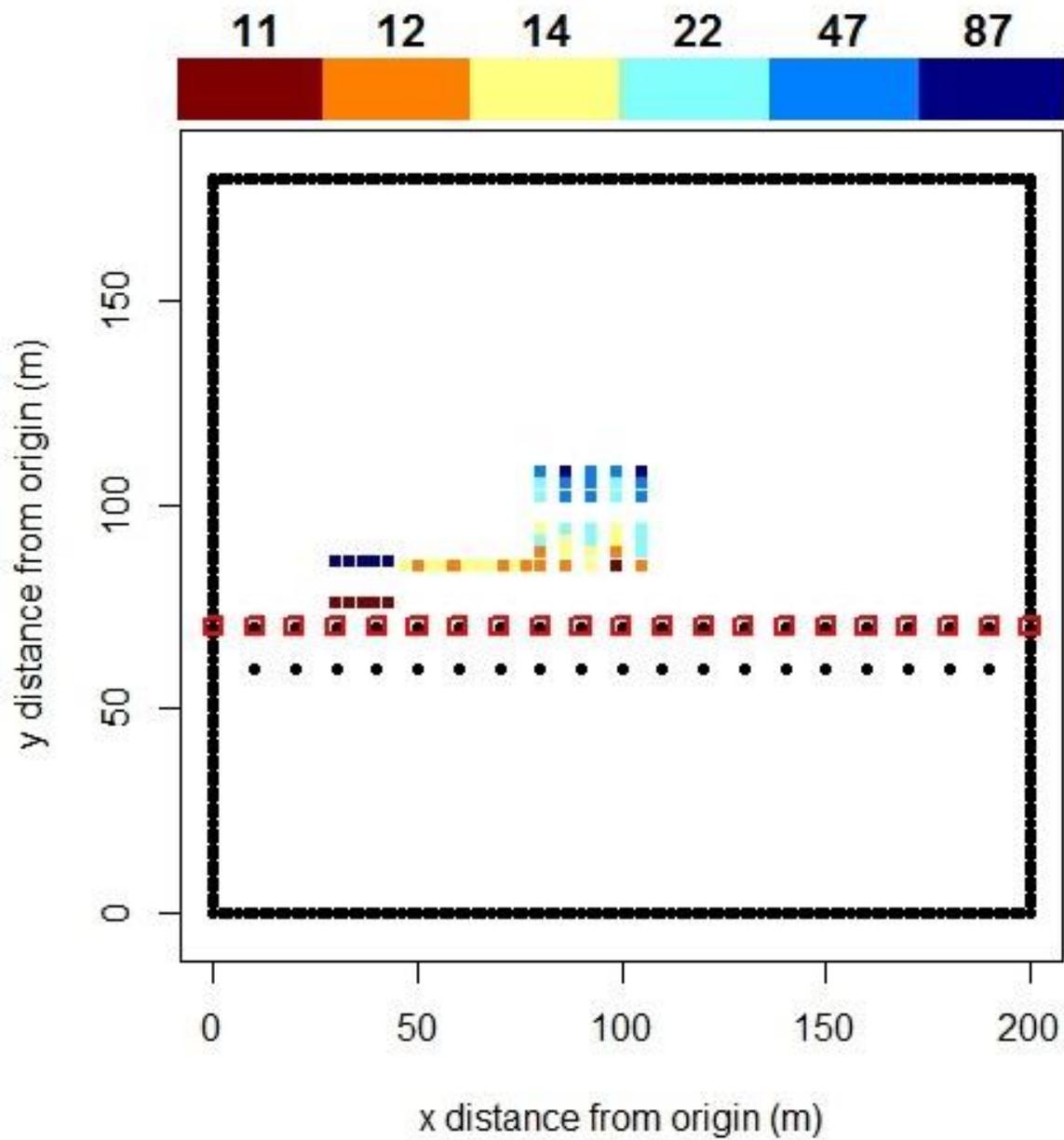
Figure S29. Bakken well pad: Cumulative distribution function of hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit point sensor for every source-receptor combination (blue = average, green = 95th percentile) and the best receptor R1M96 (orange = average, red = 95th percentile)
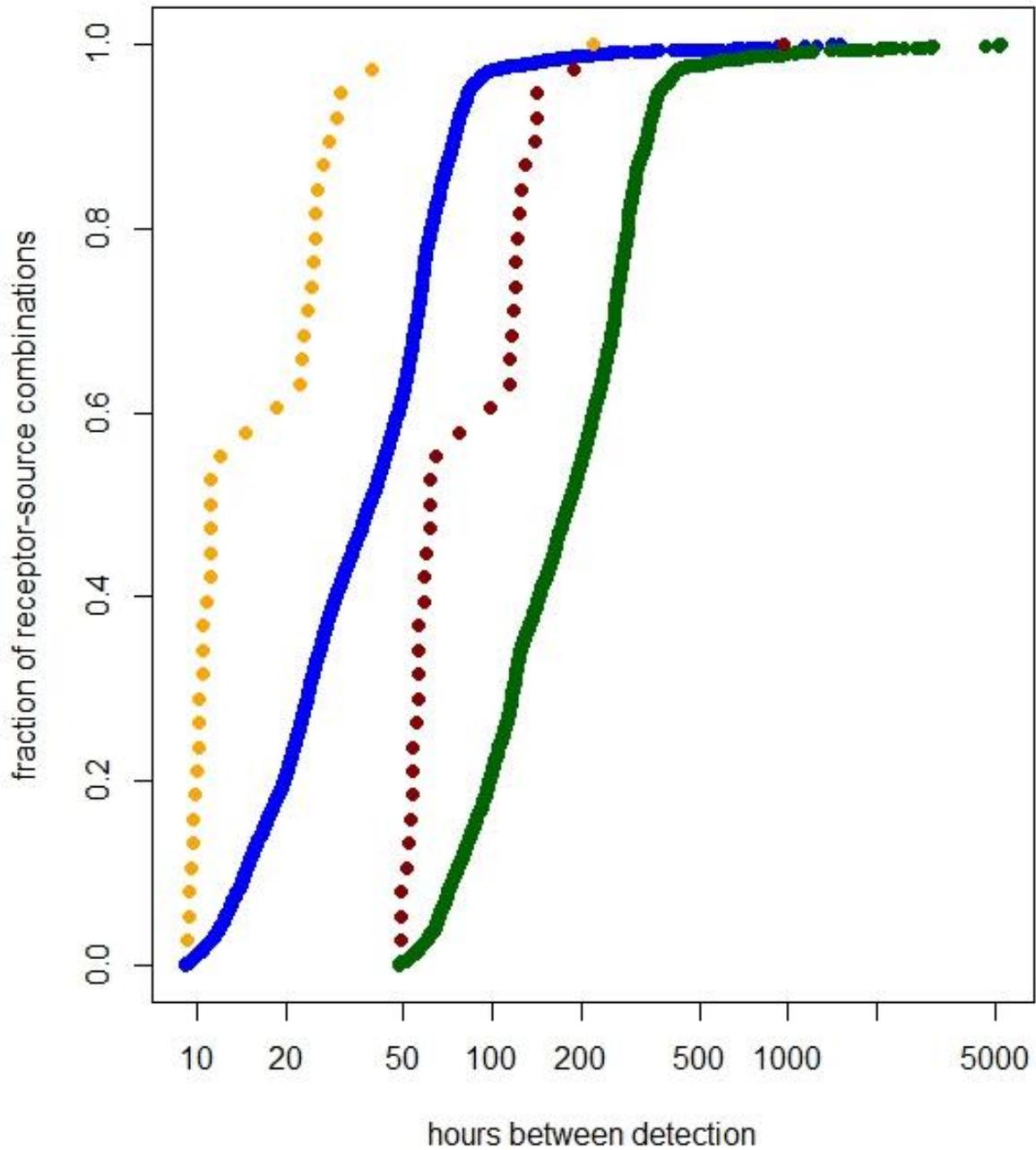
Figure S30. Bakken well pad: Site-level average hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit point sensor for every receptor at 5 m height (1st, 2.5th, 5th, 10th, 25th, 50th, & 75th percentiles). Site-level metrics are based on the median of individual sources' average hours between detection.
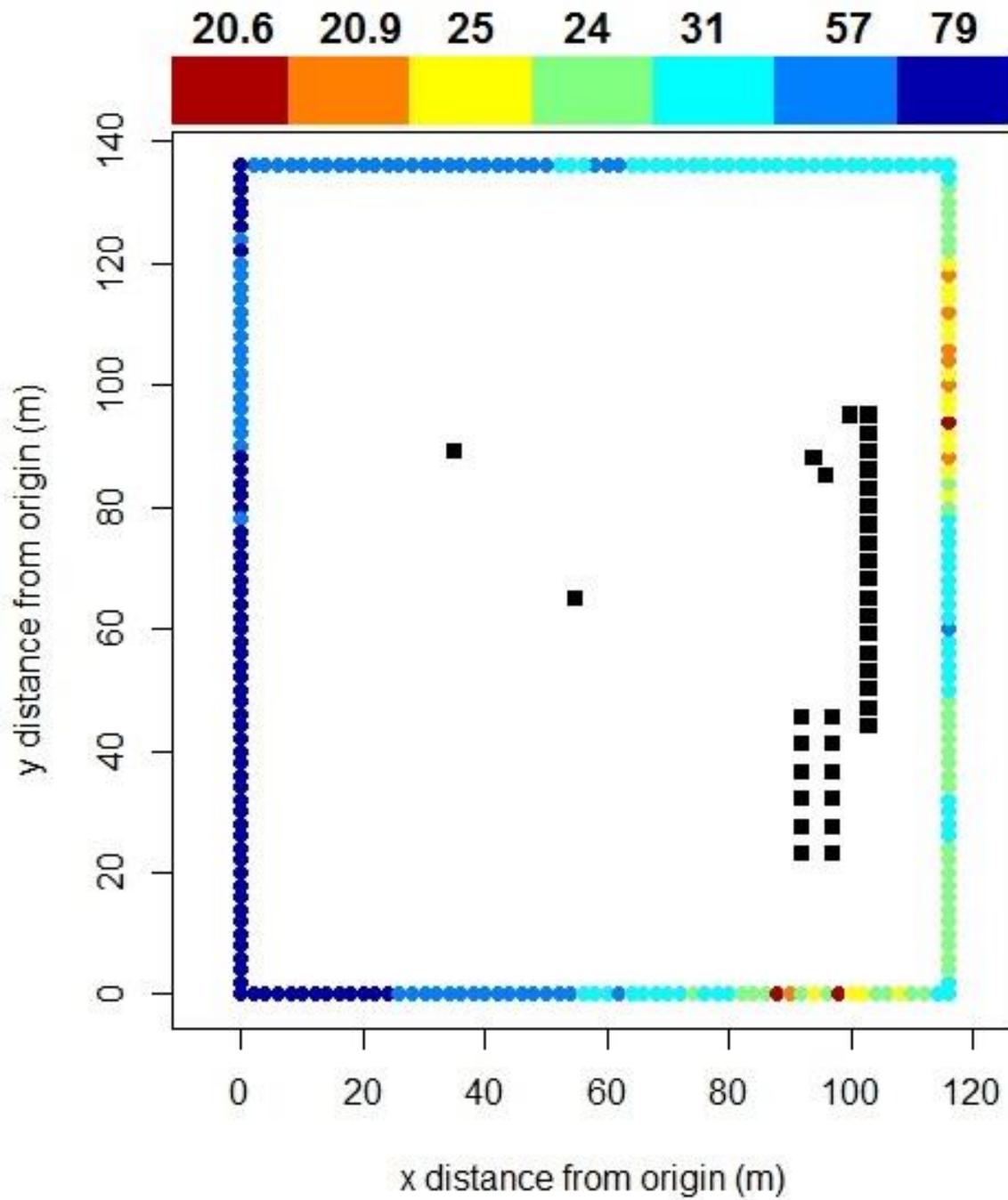
Figure S31. Bakken well pad: Cumulative distribution function of hours between detection of a 150 scfh emission rate with a 0.5 ppm detection limit open path sensor for every source-path combination (blue = average, green = 95th percentile) and the best path E1 (orange = average, red = 95th percentile)
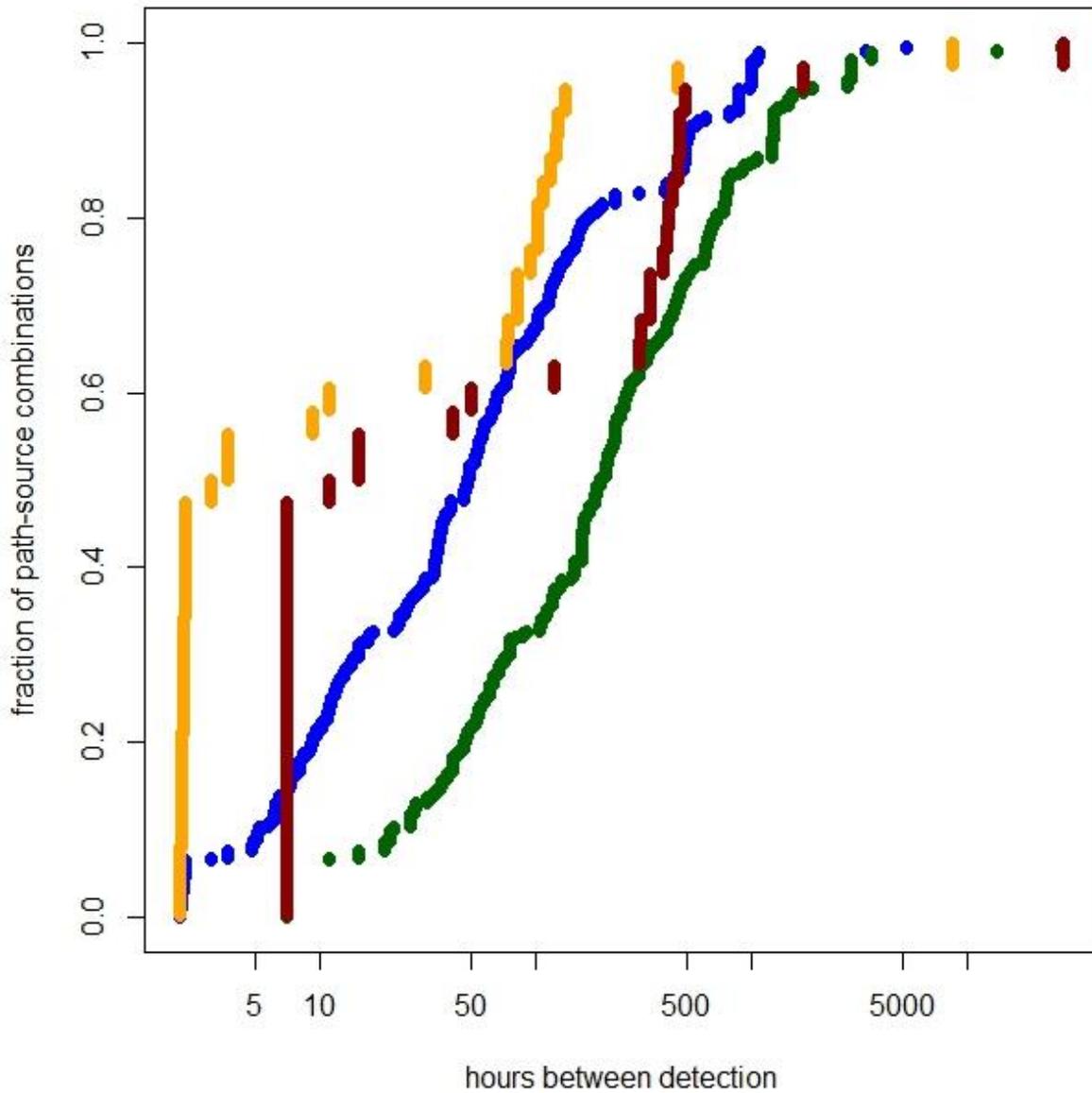
Figure S32. Bakken well pad: Average hours between detection of a 150 scfh emission rate from individual sources (10th, 25th, 50th, 75th, & 90th percentiles) with a 0.5 ppm detection limit open path sensor deployed across the best path at 5 m (E5, highlighted in red)
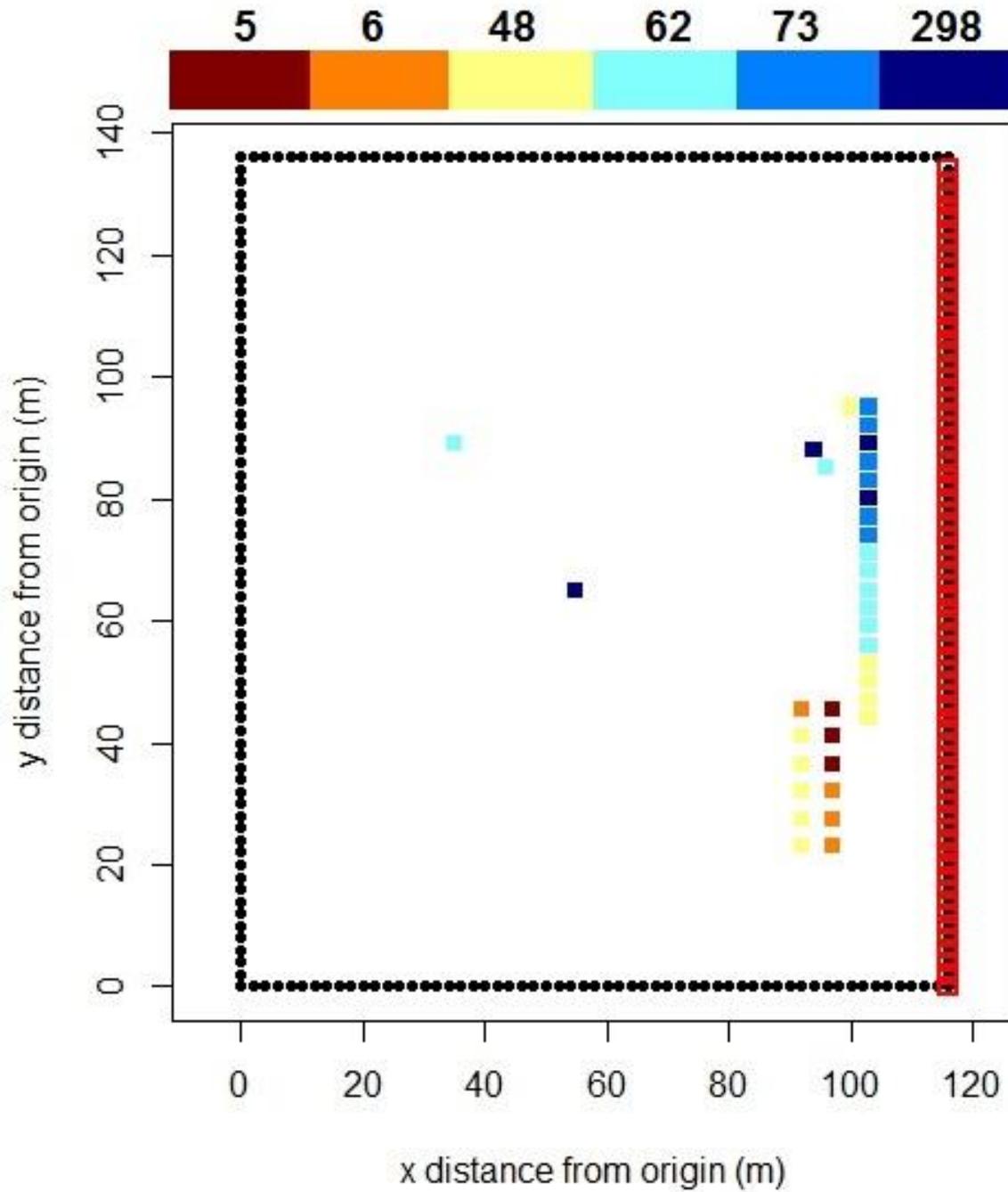
Figure S33. Bakken well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources (10th, 25th, 50th, 75th, & 90th percentiles) with a 0.1 ppm detection limit point sensor located at the receptor R1M100 (highlighted in red).This receptor had the lowest median of individual equipment leak sources' average hours between detection.
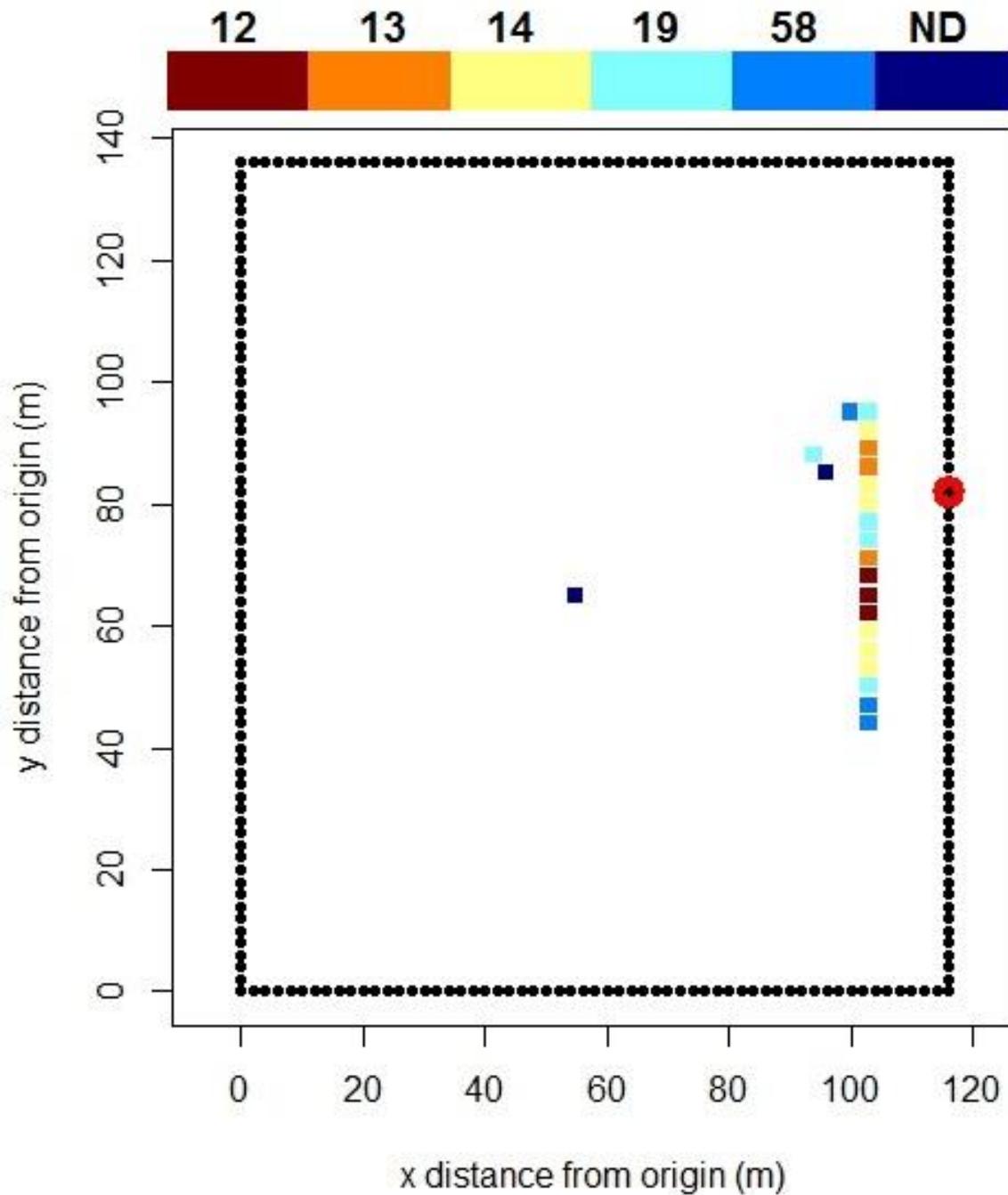
Figure S34. Bakken well pad: Average hours between detection of a 6 scfh leak from individual equipment leak sources with a 0.1 ppm detection limit open path sensor deployed across the path the lowest median of individual equipment leak sources' average hours between detection (E1, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
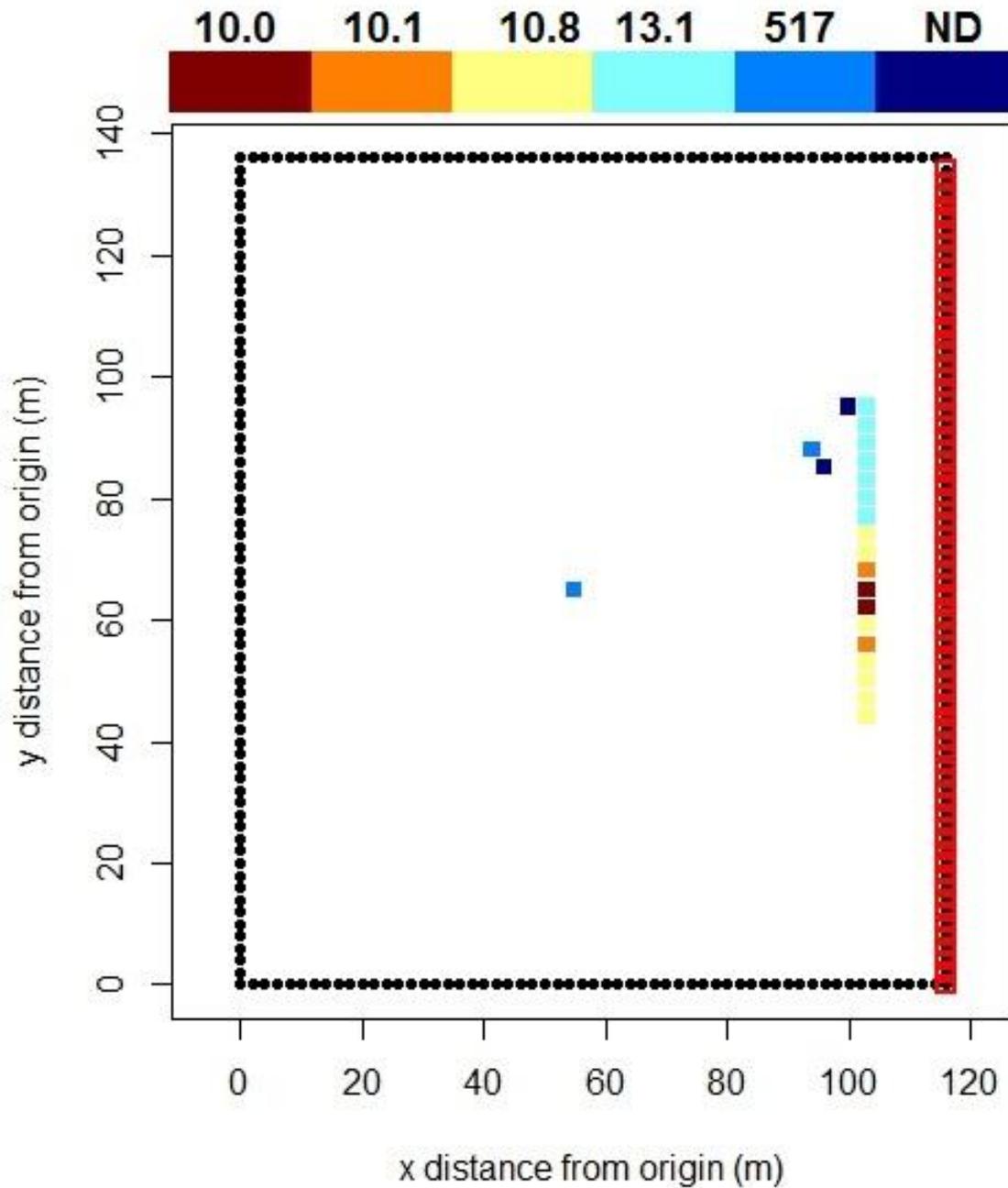
Figure S35. Bakken well pad: Average hours between enhancement ≥ 0.1 ppm from normal emissions of individual vented sources (10th, 25th, 50th, 75th, & 90th percentiles) with a sensor located at the receptor R1M251 (highlighted in red).This receptor had the lowest average hours between detection for the equipment leak source with the highest time between detection.
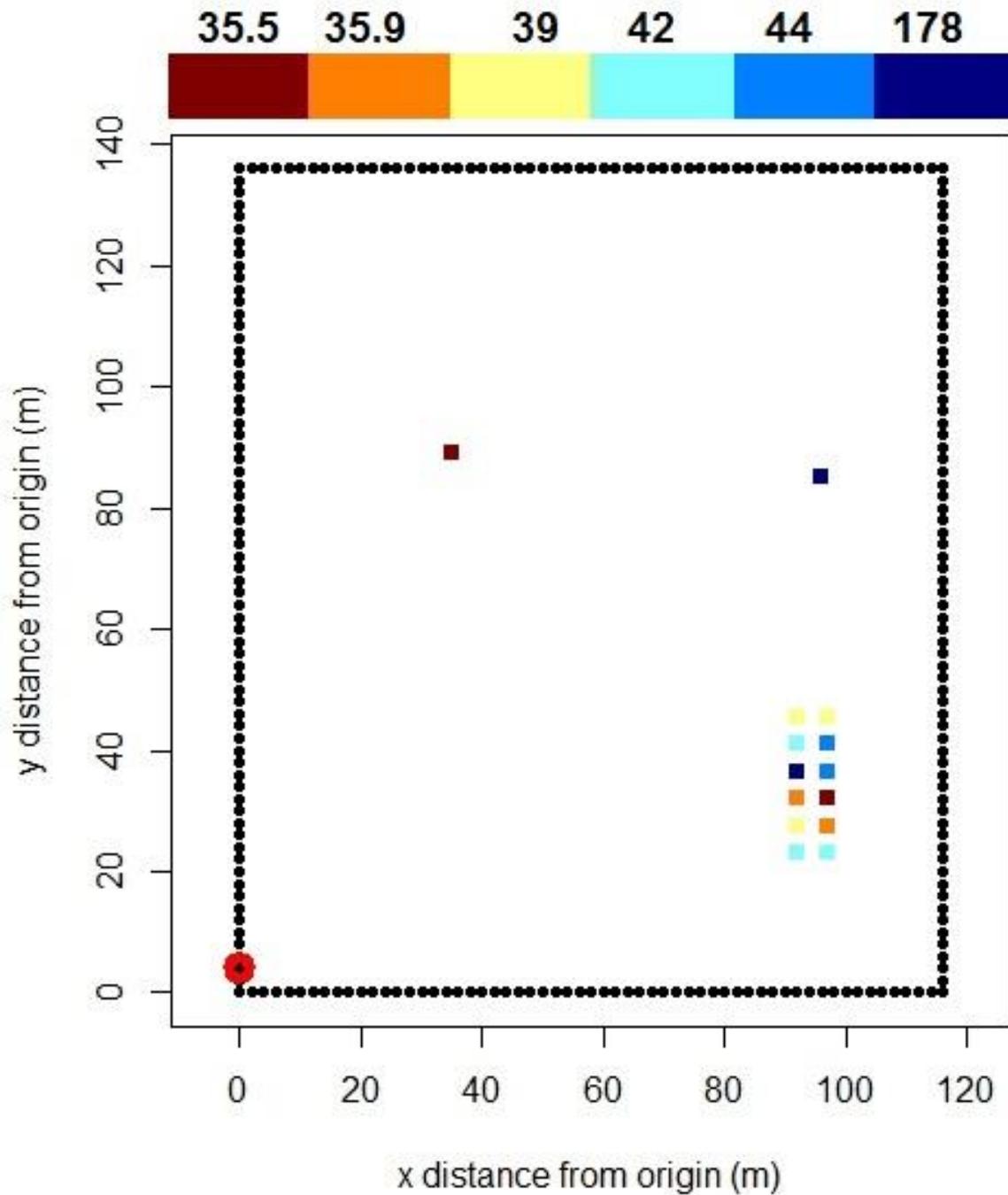
Figure S36. Bakken well pad: Average hours between enhancement ≥ 0.1 ppm from normal emissions of individual vented sources with an open path sensor deployed across path the lowest median of individual equipment leak sources' average hours between detection (E1, highlighted in red). Colors show the quantiles of time between detection (10th, 25th, 50th, 75th, & 90th percentiles).
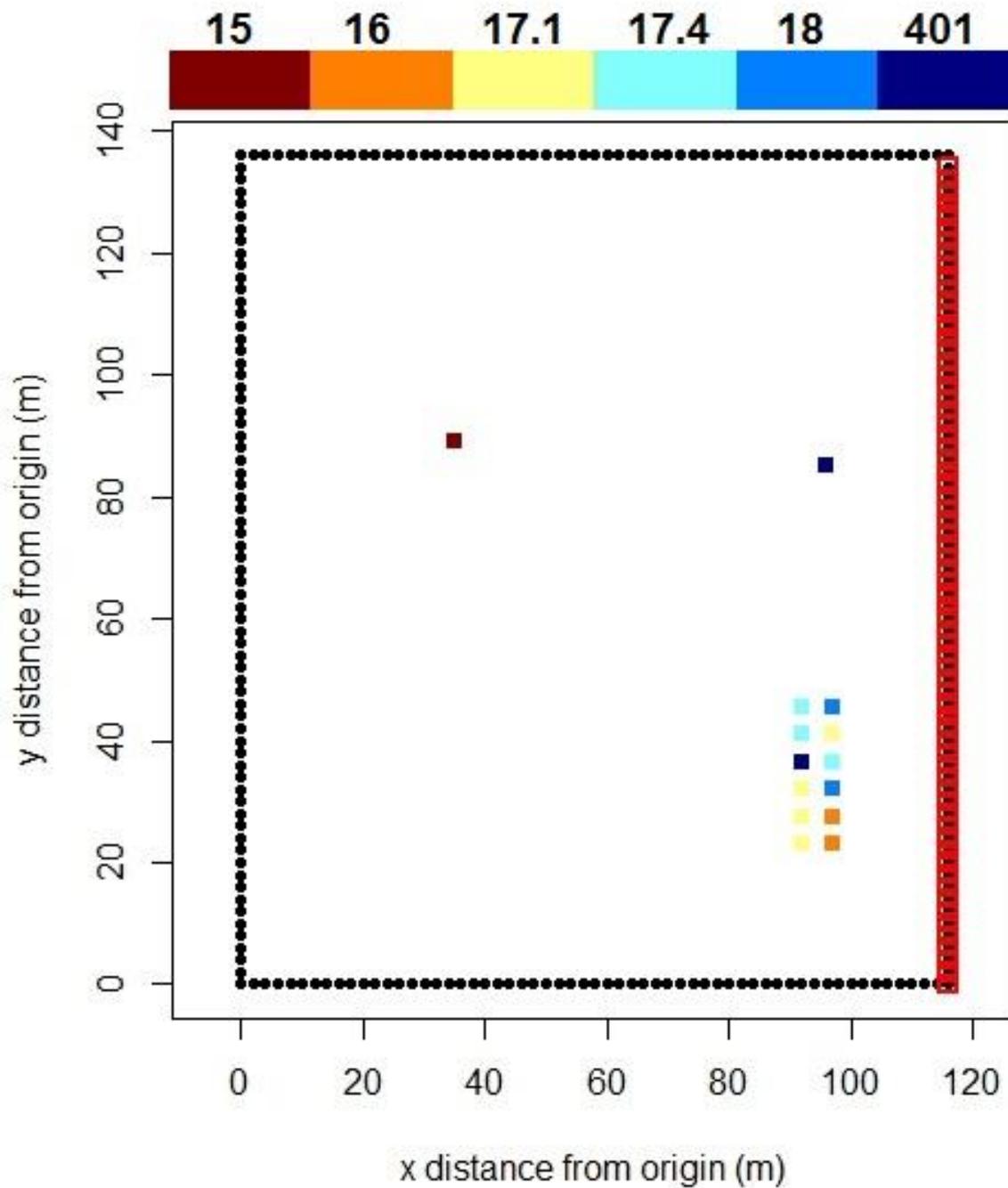
CONCLUSION

Methane super-emitters in the oil and gas (O&G) supply chain have been evaluated in three chapters that address the quantification, assessment, and mitigation of emissions. Together, these chapters demonstrate the substantial contribution of super-emitters to total supply chain emissions, the stochasticity of their occurrence, and their potential mitigation with continuous leak detection systems.

Chapter 1, *Methane emissions from the natural gas supply chain*, summarizes the state of the knowledge of O&G methane emissions. Several recent studies have measured methane emissions from facilities and components across the supply chain. Most of these studies have found highly skewed emission rate distributions with a small number of sites or components responsible for the majority of emissions. Top-down measurements of regional emissions typically have found higher emissions that bottom-up estimates, which may be due to exclusion of super-emitters in most emission inventories.

Chapter 2, *Constructing a spatially resolved methane emission inventory for the Barnett Shale Region,* integrated data from a coordinated research campaign in the Barnett Shale and a national study of gathering and processing facilities to estimate gridded, source-specific methane emissions in the Barnett Shale region of north-central Texas. Spatially-referenced activity data such as the location of individual facilities was based on multiple databases. A comprehensive count of gathering stations was higher than reported in any single source. Super-emitter emissions from well pads, small processing plants, large processing plants, and compressor stations were accounted for by developing facility emission factors with two-phase Monte Carlo simulations that blend together measurement data collected using both unbiased sampling and targeted sampling of high emission facilities. The emission inventory estimated that October 2013 emissions from the 25-county Barnett Shale region were 72,300 (63,400 – 82,400) kg CH_4 hr^{-1} from total sources and 46,200 (40,000 – 54,100) kg CH_4 hr^{-1} from O&G sources. Super-emitters, which include less than 2% of sites, are responsible for 19% of O&G emissions. Compared to alternative inventories based on the U.S. EPA Greenhouse Gas Inventory, U.S. EPA Greenhouse Gas Reporting Program, and Emissions Database for Global Atmospheric Research, the custom emission inventory was higher by factors of 1.5, 2.7, and 4.3, respectively. The higher estimate was due to the inventory utilizing both more accurate activity data and emission factors that account for super-emitters. In contrast to

several previous studies, the custom inventory's bottom-up emission estimate was not significantly different than top-down estimates of regional O&G methane emissions.

Chapter 3, *Aerial surveys of elevated hydrocarbon emissions from oil and gas production sites*, analyzed the results of a helicopter-based infrared camera survey of over 8,000 well pads in seven basins to identify sources with high hydrocarbon emissions (including methane). Based on a controlled release and aircraft-based quantification of methane emissions at a subset of sites, the survey detection limit was estimated to be $1 - 3$ g hydrocarbon s^{-1}, which is at the very high end of well pad component emissions observed in other studies. Four percent of sites were observed to have high emissions with the prevalence ranging from 1% in the Powder River to 14% in the Bakken. The occurrence of observed emissions had several statistically significant but weak correlations to site parameters such as oil production, but the most predictive multi-parameter model only explained 14% of the variance. This indicates that the occurrence of high emissions is mainly stochastic and therefore frequent monitoring of all sites is required to quickly detect super-emitters. Tank vents and hatches were the source of over 90% of observed emissions. To determine if tank flashing could explain the observed prevalence of high emissions, a tank flashing analysis was performed to estimate the number of sites expected to have emissions above the detection limit at any one time. For most basins, observed frequencies were higher than expected from tank flashing emissions if tank emission control systems had been functioning effectively. This finding was supported by the observation of tank hatch emissions at sites with control devices, which indicates control systems have poor capture efficiency. Therefore, tanks are a good target for mitigating emissions but the proper design, maintenance, and inspection of control systems is necessary to assure that emissions are reduced effectively.

Chapter 4, *Site-level Gaussian dispersion model to optimize the deployment of continuous methane sensors*, evaluated the ability of continuous leak detection systems to detect methane emissions at O&G well pads with point or open path sensors. Three example well pads in the Eagle Ford, Fayetteville, and Bakken Shale were modeled using their equipment layout and local meteorological data. Gaussian dispersion equations were used to estimate hourly concentration enhancement from individual sources at multiple receptors based on ten years of meteorological data. The model calculated the average and 95th

percentile detection time of two emission rates (6 and 150 standard cubic feet per hour) using sensors with three detection limits (0.5, 0.1, and 1 ppm CH_4 enhancement above background). At all three sites, there were optimum locations for both point and open path sensors with a 0.5 ppm detection limit to detect 150 scfh leaks within a few days. This performance level is consistent with the initial goals of the Methane Detectors Challenge, which will deploy first generation leak detection systems at pilot sites in 2016. The detection of 6 scfh leaks was poor for 0.5 ppm detection limit sensors but substantially improved for 0.1 ppm detection limit sensors. Accordingly, first generation continuous leak detection systems are expected to quickly detect super-emitters. Modest improvements to their detection limit will allow systems to also rapidly detect smaller equipment leaks. Continuous leak detection systems are a promising emerging technology for cost-effectively mitigating O&G methane emissions.

Outlook

Super-emitters are an important contributor to O&G methane emissions and therefore their rapid identification and mitigation is critical for reducing the climate impact of natural gas. The discrepancy between top-down and bottom-up emission estimates reported by previous studies appears to be partially caused by the exclusion of super-emitters from emission inventories. As a follow-up to the work in Chapter 1, I co-authored a manuscript that revised the Barnett Shale gridded inventory with an improved approach to account for super-emitters.[1] Rather than using two-phase Monte Carlo simulations to draw from two separate distributions with set super-emitter probabilities, the new approach combines unbiased and high-biased measurements into single log-normal distributions with statistical estimators. The updated O&G emission estimate is 28% higher than the estimate presented in Chapter 1 and within 10% of the top-down estimate. The difference from the earlier estimate is mainly due to a higher production site emission factor with even greater contribution from super-emitters – the top 2% of sites were estimated to be responsible for 50% of emissions. Ongoing research is evaluating what fraction of production site emissions can be attributed to component-level emissions. It is likely that a substantial portion of the emissions from the highest emitting sites cannot be explained by normally functioning components. Based on the insights of Chapter 2, most of these super-emitter sites probably have abnormally high tank emissions, which could be caused by separator malfunctions or ineffective tank

control systems. Future research should quantify methane emissions from tank super-emitters in basins where top-down emission estimates exceed bottom-up estimates to determine if these sources account for the gap.

Future regulations and voluntary initiatives to reduce O&G methane emissions must account for super-emitters and the stochasticity of their occurrence. Although super-emitters cannot be predicted, there likely are practices that can reduce their frequency. In particular, proper design and maintenance of facilities including control device could limit malfunctions that lead to abnormally high emissions. Once super-emitters do occur, their rapid detection and repair is crucial for minimizing emissions. Similar to many large industrial facilities, O&G facilities should utilize process monitoring to continuously assess equipment status so that operators are immediately aware of malfunctions or ongoing issues likely to cause future equipment failure. Many production sites already have systems that measure oil and gas production rates and transmit the data to operators. These systems could be upgraded to monitor and transmit information from equipment often associated with high emissions, such as contact sensors that indicate when a tank hatch is open. Finally, continuous leak detection systems are emerging as a cost-effective tool for quickly identifying the occurrence of methane emissions, particularly from large leaks. These steps can be a part of comprehensive efforts to reduce methane emissions from both super-emitters and widespread, smaller sources in the O&G supply chain to minimize climate impacts.

CONCLUSION REFERENCES

Zavala-Araiza, D., Lyon, D.R., Alvarez, R.A., Davis, K.J., Harriss, R., Herndon, S.C., Karion, A., Kort, E.A., Lamb, B.K., Lan, X. and Marchese, A.J., 2015. Reconciling divergent estimates of oil and gas methane emissions. *Proceedings of the National Academy of Sciences*, 112(51), pp.15597-15602.