

DISSERTATION

MODELING METHANE EMISSIONS FROM US NATURAL GAS OPERATIONS: NATIONAL
GATHERING STATION EMISSION FACTOR DEVELOPMENT AND
FACILITY/REGIONAL-SCALE TOP-DOWN TO BOTTOM-UP RECONCILIATIONS

Submitted by

Timothy L. Vaughn

Department of Mechanical Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2017

Doctoral Committee:

Advisor: Anthony J. Marchese

Co-Advisor: Azer P. Yalin

Daniel B. Olsen

Jean D. Opsomer

Copyright by Timothy L. Vaughn 2017

All Rights Reserved

ABSTRACT

MODELING METHANE EMISSIONS FROM US NATURAL GAS OPERATIONS: NATIONAL GATHERING STATION EMISSION FACTOR DEVELOPMENT AND FACILITY/REGIONAL-SCALE TOP-DOWN TO BOTTOM-UP RECONCILIATIONS

United States natural gas dry production increased by 47% between 2005 and 2015 due to the widespread use of horizontal drilling and hydraulic fracturing to extract gas from shale and other tight formations. Natural gas production and consumption is projected to continue to increase for the foreseeable future. In 2016, the natural gas supply chain delivered 29% of the energy used in the U.S., and natural gas surpassed coal as the leading electricity generating source for the first time in U.S. history.

When combusted, natural gas produces less CO₂ per unit energy released compared to coal or petroleum. However, uncombusted methane (the primary component of natural gas) has a global warming potential 30 times higher than CO₂ on a 100 year time horizon (including oxidation to CO₂, but excluding climate-carbon feedbacks). Therefore, the net greenhouse gas impacts resulting from displacement of coal and petroleum by natural gas depend on the emission rate of uncombusted natural gas. Short term climate benefits resulting from coal substitution, for example, are lost if the net rate of methane (CH₄) emission from the natural gas supply chain exceeds 3–4%.

Three studies were conducted to quantify CH₄ emissions from the natural gas industry. In particular, these studies focused on quantifying emissions from the gathering and processing sector and reconciling emissions estimates developed using top-down (tracer flux and aircraft) vs. bottom-up (on-site component-level) measurement approaches.

In the first study, facility-level CH₄ emissions measurements were made at 114 natural gas gathering facilities and 16 processing plants in 13 U.S. states during a 20-week field campaign conducted from October 2013 through April 2014. Measurement results were combined with

facility counts obtained from state air permit databases and national inventories in a Monte Carlo simulation to estimate CH₄ emissions from U.S. natural gas gathering and processing operations. Annual CH₄ emissions from normal operations at gathering facilities totaled 1699 Gg (95% CI=1539–1863 Gg), while normal operations at processing plants totaled 505 Gg (95% CI=459–548 Gg). CH₄ emissions from abnormal operations at gathering facilities were estimated in a separate Monte Carlo simulation based on field observations and a sub-set of field measurements. These emissions totaled 169 Gg (+426%/-96%).

In the second study, coordinated dual-tracer, aircraft-based, and direct component-level measurements were made at midstream natural gas gathering and boosting stations in the Fayetteville shale in Arkansas, USA. On-site component-level measurements were combined with engineering estimates to generate comprehensive facility-level CH₄ emission rate estimates (“study on-site estimates (SOE)”) comparable to tracer and aircraft measurements. Concurrent measurements at 14 normally-operating facilities showed a strong correlation between tracer and SOE, but indicated that tracer measurements estimated lower emissions (regression of tracer to SOE=0.91 (95% CI=0.83–0.99, R²=0.89). Tracer and SOE 95% confidence intervals overlapped at 11/14 facilities. Contemporaneous measurements at six facilities suggested that aircraft measurements estimated higher emissions than SOE. Aircraft and study on-site estimate 95% confidence intervals overlapped at 3/6 facilities.

In the third study, a detailed spatiotemporal inventory model was developed and used to reconcile top-down and bottom-up CH₄ emission estimates from natural gas infrastructure and other sources in the Fayetteville shale on two consecutive days. On Thursday October 1, 2015 13:00–15:00 CDT top-down aircraft mass balance flights estimated 28.7 (20.1–37.3 Mg/h 95% CI) from the study area, while the bottom-up ground level area estimate predicted 23.9 (20.9–27.3 Mg/h 95% CI). On Friday October 2, 2015 14:30–16:30 CDT top-down estimated 36.7 (21.3–52.1 Mg/h 95% CI), while bottom-up estimated 21.1 (18.4–24.2 Mg/h 95% CI). Production and gathering activities were the largest contributors to modeled CH₄ emissions. In contrast to prior studies, comparisons on two consecutive days

indicated overlapping confidence intervals between top-down aircraft estimates and bottom-up inventory-driven estimates. Operator participation and extensive activity data proved critical in understanding emissions as observed by aircraft. In particular, the agreement obtained was possible only because bottom-up models included the variability in production maintenance activities, which showed substantially higher emissions during daytime hours when aircraft-based measurements were performed. Results indicated that that poor activity estimates (counts and timing) for large episodic events likely drives divergence in CH₄ emission estimates from production basins, and that even more precise activity data would be required to improve agreement between these two approaches.

ACKNOWLEDGEMENTS

I'd like to thank my advisor, Prof. Anthony Marchese, for providing me with many opportunities to work on interesting and challenging projects over the years. I would also like to thank my committee members: Prof. Daniel Olsen, Prof. Azer Yalin, and Prof. Jean Opsomer for their time and expert advice, along with Dan Zimmerle.

None of the work presented here could have been successful without the dedication, hard-work, and cooperation of a large number of people. It was a new and exciting experience to participate in projects of this scale. I'd like to thank all involved, but especially the great folks I had a chance to work with directly.

The work presented in Chapter 2 ("G&P") was enabled by financial and in-kind support from the Environmental Defense Fund (EDF), Access Midstream, Anadarko Petroleum Corporation, Hess Corporation, Southwestern Energy, Williams Companies, and DCP Midstream. For the G&P project, my hat goes off to: Prof. Marchese, Dan Zimmerle, Laurie Williams, Dave Martinez, and Jerry Duggan at CSU/Fort Lewis; Scott, Tara, and Rob at Aerodyne; and Austin, Daniel, and Melissa at CMU. I'd also like to thank all of the partner company liaisons, and field operations personnel for their patience, hard work and support during the field campaign.

The work presented in Chapters 3 and 4 ("RPSEA") was enabled by financial and in-kind support from the U.S. Department of Energy RPSEA/NETL contract no. 12122-95/DE-AC26-07NT426777, the Colorado Energy Research Collaboratory, the National Oceanic and Atmospheric Administration, Southwestern Energy, XTO Energy, a subsidiary of Exxon-Mobil, Chevron, Statoil, the American Gas Association, CenterPoint, Enable Midstream Partners, Kinder Morgan, and BHP Billiton. For the RPSEA project, I'd like to thank: Clay Bell for playing a key role in the field campaign and helpful modeling discussions, along with Dan Zimmerle, Cody Pickering, and Laurie Williams at CSU/Fort Lewis; Tara, Scott, and Rob at Aerodyne; Gaby and Stefan at NOAA/UC CIRES; Steve at Scientific Aviation;

the DGC LDAR team; the AECOM team; and all of the other measurements teams—UWY, the NOAA van team, GHD, and SWN LDAR. I'd also like to thank all of the field operators, and their supervisors, for their patience, hard work and support throughout the field campaign. Members of the various committees deserve thanks as well: Dag, Garvin, Russ, Doug, Gib, Tecle, Desikan, Fiji, and Pam.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDEGMENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER 1: Introduction	1
1.1 Overview of Dissertation	4
CHAPTER 2: Estimation of Methane Emissions from the US Natural Gas Gathering and Processing Sector	6
2.1 Introduction.	6
2.2 Gathering Model.	9
2.2.1 State Datasets	10
2.2.2 Estimation of Total State Gathering Facility Counts.	11
2.2.3 Estimation of Confidence Intervals for Total State Gathering Facility Counts	13
2.2.4 Facility Natural Gas Throughput vs Compressor Horsepower	15
2.2.5 Facility-level Emission Rates vs Natural Gas Throughput	16
2.2.6 Gathering Model Walk-through	19
2.3 Processing Model	22
2.3.1 Processing Plant Dataset	23
2.3.2 Processing Plant Model Walk-through.	26
2.4 Monte Carlo Model Results	27
2.5 Conclusions	30
CHAPTER 3: Comparing Facility-Level Methane Emission Rate Estimates at Natural Gas Gathering and Boosting Stations	33
3.1 Introduction.	33
3.2 Field Study	35
3.3 Measurements	37
3.3.1 Tracer Measurements	42
3.3.2 Aircraft Measurements	43
3.3.3 On-site Measurements	43
3.3.4 Stack Test Measurements	46
3.4 Model Description	47
3.5 Results	53
3.5.1 Tracer Facility Estimate and Study On-site Estimate Comparison	56
3.5.2 Aircraft Facility Estimate and Study On-site Estimate Comparison	59
3.6 Conclusions	64

CHAPTER 4:	Reconciling Top-Down and Bottom-Up Methane Emission Estimates from Natural Gas Operations in the Fayetteville Shale	67
4.1	Introduction	67
4.2	Methods: Model Description	70
4.3	Oil and Gas Related Methane Sources	71
4.3.1	Production	71
4.3.2	Gathering	73
4.3.3	Transmission	81
4.3.4	Distribution	82
4.4	Non-Oil and Gas Related Methane Sources	85
4.4.1	Livestock	85
4.4.2	Rice Cultivation	88
4.4.3	Wetlands	89
4.4.4	Geologic Seeps	92
4.4.5	Landfills	93
4.4.6	Wastewater Treatment	94
4.4.7	GHGRP Facilities	96
4.5	Results and Discussion	97
4.5.1	Temporally Varying Ground-level Area Estimate (GLAE)	97
4.5.2	Comparisons Considering Flight Time	102
4.6	Conclusions	108
CHAPTER 5:	Conclusions and Recommendations	110
REFERENCES	113
APPENDIX A:	Estimation of Methane Emissions from the U.S. Natural Gas Gathering and Processing Sector	121
APPENDIX B:	Comparing Facility-Level Methane Emission Rate Estimates at Natural Gas Gathering and Boosting Stations	125
B.1	Alternate Method Comparisons Using SOEs Developed from Measured Dehydrator Regenerator Vents	126
B.1.1	SOE and Overall Results Summary	126
B.1.2	Tracer Facility Estimate and Study On-site Estimate Comparison	128
B.1.3	Aircraft Facility Estimate and Study On-site Estimate Comparison	130
B.2	Variance-Weighted Least-Squares Regression	131
APPENDIX C:	Reconciling Top-Down and Bottom-Up Methane Emission Estimates from Natural Gas Operations in the Fayetteville Shale	134
C.1	Ground-level Area Estimate (GLAE) Results Corresponding Aircraft Mass Balance Flights	135
C.1.1	Flight Window GLAE Results: October 1 st , 2015	135
C.1.2	Flight Window GLAE Results: October 2 nd , 2015	136

C.2	Ground-level Area Estimate (GLAE) Sensitivity Studies Corresponding to the October 1 st , 2015 Aircraft Mass Balance Flight	137
C.2.1	Scenario 1: Manual Liquid Unloading Emission Rates Increased 37%.	137
C.2.2	Scenario 2: Gathering Station Emission Rates Increased 89%	138
C.2.3	Scenario 3: Livestock Emission Rates Increased 4x	139
C.2.4	Scenario 4: Wetland Emission Rates Increased 8.5x	140
C.2.5	Scenario 5: Geologic Seep Emission Rates Increased 6.5x.	141
C.2.6	Scenario 6: Timing of Short-Duration, High-emission Rate Events	142
C.3	Hourly Ground-level Area Estimate (GLAE) Results for October 1 st and 2 nd , 2015	145
C.3.1	Hourly GLAE Results: October 1 st , 2015	145
C.3.2	Hourly GLAE Results: October 2 nd , 2015	169
LIST OF ABBREVIATIONS		193

LIST OF TABLES

2.1	Summary data table of gathering facility counts for eight modeled states	12
2.2	Gathering facility counts by state with confidence intervals	14
2.3	Monte Carlo simulation results for modeled states	29
3.1	Measurement team availability during the field campaign	37
3.2	Measured facility and method comparison counts	38
3.3	On-site direct measurement (ODM) counts by source category with exceptions .	44
3.4	Dehydrator still vent emission factor comparison	45
3.5	Combustion slip emission factor comparison	47
3.6	Cumulative emissions: SOE to tracer comparison	59
3.7	Cumulative emissions: SOE to aircraft comparison	62
4.1	Ground level area estimate (GLAE) model source categories summary table . .	71
4.2	Production source categories summary table	73
4.3	Gathering source categories summary table	74
4.4	Gathering high-flow measurement categories summary table	75
4.5	Distribution source categories summary table	83
4.6	Distribution sector measurement table	84
4.7	USDA AR county level livestock census data	86
4.8	Modeled USDA AR county level livestock census data	87
4.9	EPA GHGI and IPCC livestock methane emission factors	87
4.10	Rice cultivation emission factor used in model	88
4.11	Wetland emission factors used in model	90
4.12	Geologic seep emission factors used in model	92
4.13	Landfill emission factors used in model	94
4.14	Wastewater activity data used in model	95
4.15	Wastewater emission factors used in model	95
4.16	GHGRP facility-level emission factors used in model	96
4.17	Ground-level area estimate to national inventory comparison	101

LIST OF FIGURES

2.1	Gathering and processing Monte Carlo model overview	8
2.2	Facility natural gas throughput vs operating compressor engine horsepower . . .	15
2.3	Facility natural gas throughput vs installed compressor engine horsepower . . .	16
2.4	Facility-level methane (CH ₄) emission rates vs throughput	17
2.5	Throughput normalized facility-level emission rates vs throughput	19
2.6	Nearest-neighbors modeling approach	22
2.7	Daily average natural gas throughput at U.S. processing plants	25
2.8	Effect of number of nearest-neighbors on CH ₄ emissions predicted in processing plant model	26
2.9	Monte Carlo simulation results for modeled states and processing plants	28
2.10	Processing plant model result CDF	30
3.1	Typical CD gathering station in Fayetteville shale	34
3.2	Gathering stations within study area	36
3.3	Facility total compressor engine horsepower comparison	37
3.4	Gathering station 96 was excluded from comparisons	41
3.5	Gathering station 98 was excluded from comparisons	42
3.6	Facility total operating horsepower vs throughput	45
3.7	Measured combustion slip	46
3.8	High-flow over-range example	50
3.9	All comparable facility level emission rate estimates	55
3.10	Tracer facility estimate and study on-site estimate difference plot	56
3.11	Tracer facility estimate and study on-site estimate variance-weighted least-squares regression	58
3.12	Aircraft facility estimate and study on-site estimate difference plot	60
3.13	Aircraft facility estimate and study on-site estimate variance-weighted least-squares regression	61
3.14	GHGI Comparison CDF	64
4.1	Sub-basin study area overview	68
4.2	Spatial boundary of the study area that contributes to the ground level area estimate	70
4.3	Fayetteville production well pads	72
4.4	Fayetteville gathering stations	73
4.5	Fayetteville transmission stations	81
4.6	Locations served by local distribution	82
4.7	Livestock data counties	85
4.8	Wetlands within the study area	90
4.9	Landfills within the study area	93
4.10	Hourly study area emission rate estimates (GLAEs) for October 1 st and 2 nd . .	97
4.11	Study area methane emission estimates during aircraft mass balance flights . . .	102
4.12	Downwind transect estimate infographic	103

4.13	Downwind transect estimates during aircraft mass balance flights	104
4.14	Manual liquid unloading and gathering station emission rate sensitivity analyses	105
4.15	Downwind transect estimate capturing compressor engine start-up	108

CHAPTER 1

INTRODUCTION

United States natural gas dry production has increased steadily from 18,050,598 MMcf (million cubic feet) in 2005 to 26,459,310 MMcf in 2015 (a 47% increase), before decreasing slightly in 2016¹. This increase in production has resulted from widespread use of horizontal drilling and hydraulic fracturing to extract gas from shale and other tight formations¹. U.S. Energy Information Agency (EIA) forecasts predict that natural gas production and consumption will continue to increase for the foreseeable future². Natural gas now makes up the largest share of U.S. electrical generating capacity and was the leading generation source in 2016³, surpassing coal for the first time in U.S. history. Increased natural gas production and usage are driving a need to understand overall system loss rates from the U.S. natural gas supply chain.

Natural gas produces less carbon dioxide (CO_2) when combusted than coal or petroleum on a per unit energy basis, and is often suggested as a bridge fuel to a lower-carbon energy sector. However, total greenhouse gas impacts from natural gas use are highly dependent on the emission rate of un-combusted natural gas⁴ because methane (CH_4), the primary component of natural gas, has a global warming potential 30 times higher than CO_2 on a 100 year time horizon (including oxidation to CO_2 , but excluding climate-carbon feedbacks)⁵. These factors highlight the importance of reducing lifecycle methane emissions, since short term climate benefits from coal substitution are lost if net methane emission rates exceed 3–4%⁴.

Atmospheric methane enhancement has increased since 2007⁶, concomitant with the aforementioned increase in natural gas production. However, the cause of atmospheric methane enhancement is poorly understood⁷ and source attribution is undetermined⁶. The 2016 Environmental Protection Agency (EPA) greenhouse gas inventory (GHGI)⁸ indicates that natural gas systems are currently the leading methane emission source in the U.S..

Efforts to accurately quantify methane emissions from natural gas systems are currently underway. These efforts aim to identify targets for emission reduction to reap potential climate benefits, reduce potential impacts on local air quality, improve safety, and minimize lost product.

Broadly speaking, the natural gas supply chain is divided into the following sectors: production, gathering and processing, transmission and storage, and distribution. The production sector includes completed wells producing natural gas, and associated equipment at the well pad such as tanks and separators. The gathering and processing sector includes gathering pipelines, gathering stations, and processing plants. Generally, these “midstream” facilities collect gas from production wells, condition it to pipeline quality, and transfer it to transmission and storage, or distribution. Conditioning may entail oil or condensate removal, water removal, natural gas liquid separation, acid gas removal, compression, fractionation, etc. The transmission and storage sector includes transmission pipelines, transmission compressor stations, and gas storage facilities. This sector transports large quantities of natural gas from producing regions to consuming regions through large-diameter interstate pipelines operating at high pressures. Some portion of transported gas may be delivered to storage facilities where it is re-injected into underground reservoirs, and withdrawn to buffer supply during periods of high demand. Distribution includes the pipelines, regulators, meters, and associated equipment used by local utilities to deliver natural gas to end users at commercial and residential locations.

The United States has a vast infrastructure supporting the production, gathering, processing, transmission, storage, and distribution of natural gas. National activity data (counts) in the 2016 U.S. EPA GHGI⁹ indicate that there are 456,140 natural gas wells, 4,999 gathering compressor stations, 431,051 miles of gathering pipeline, 668 processing plants, 301,748 miles of transmission pipeline, 1,834 transmission compressor stations, 356 storage compressor stations, and 2.2 million miles¹⁰ of local utility distribution pipelines. This infrastructure

delivered 29% of the energy used in the U.S. in 2016¹¹. This vast infrastructure makes quantification of leaks a challenging venture.

Past studies have aimed to quantify emissions from the natural gas system; the GHGI relies heavily on data from studies performed in the 1990's¹². More recent studies have developed estimates of methane emissions from production,¹³ gathering and processing,¹⁴⁻¹⁶ transmission and storage,^{17,18} and distribution¹⁹ using a combination of field measurements and modeling.

A number of techniques and technologies have been employed to measure methane emissions from natural gas facilities. Direct measurements of leaking components (pneumatic devices, connectors, flanges, etc.) may be made using high-flow samplers, calibrated bags, anemometers, flow meters, etc. Process vent emissions and compressor engine exhaust emissions can also be measured directly, but require more sophisticated methods and instrumentation. Alternatively, methane emissions from an entire facility (such as a gathering station) can be estimated from atmospheric concentration enhancements measured downwind of the facility. Many variations on this approach exist, including tracer flux,^{16,20} Other Test Method (OTM) 33a,²¹ and others. Total methane emissions from a facility²² or region²³ can also be measured via aircraft using a variety of techniques. Methane emissions have also been estimated using satellite data.²⁴

The aforementioned techniques can be employed in either “bottom-up” or “top-down” approaches to quantify leaks from natural gas facilities, or producing regions. In bottom-up estimates, emissions from individual components are measured directly, and all measurement results are summed to produce a facility-level estimate. In top-down estimates, emissions from all sources at the facility are deduced from downwind measurements of atmospheric concentration enhancements. The exact meaning of these terms varies with context. For example, the sum of individual facility-level emissions in a gas producing region may be referred to as a “bottom-up” estimate, when compared to “top-down” measurements of the

region made from aircraft or satellites, even if the individual facility-level emissions summed were the result of downwind measurements.

Accurately estimating total methane emissions from a given facility, sector, or region is a challenging endeavor. Direct measurement of all sources is possible in theory, but challenging in practice due to the large number of sources, and the variety of sources present. For example, measurement of all sources at a mid-sized gathering station would require several groups of measurement personnel with specialized skills and instrumentation, and would take several days. Emissions from certain sources vary in time, and spot measurements may not capture this variation completely. Certain sources are challenging or impossible to measure directly due to safety concerns. Individual measurements are subject to uncertainty, and obtaining sufficient measurement samples to characterize overall populations is time and cost prohibitive at the component, facility, and regional level. For these reasons, statistical techniques are often used to extrapolate measured samples to larger populations. Herein, Monte Carlo simulations are used to scale component-level measurements and activity data to aggregate facility and regional-level emissions estimates.

1.1 OVERVIEW OF DISSERTATION

This dissertation includes three studies that aimed to quantify CH₄ emissions from the natural gas industry. The first and second studies focused on the gathering and processing sector. The second and third studies focused on reconciling emissions estimates developed using top-down and bottom-up approaches.

In the first study, which is described in Chapter 2, a national estimate of methane emissions from the gathering and processing sector was developed. Measurement results from a 20-week national field campaign provided facility-level emission rates for 114 gathering facilities and 16 processing plants. Gathering facility counts and installed compressor engine horsepower were compiled by analyzing air permit data from eight U.S. states. Processing plant counts and annual natural gas throughput were compiled from existing inventories.

Activity data and emission rates were combined in a Monte Carlo model that employed a nearest-neighbor kernel smoother.

In the second study, described in Chapter 3, facility-level emission rates from tracer flux measurements and a new aircraft measurement technique were compared to study on-site estimates at natural gas gathering facilities in Arkansas. Tracer and study on-site estimates were compared at 14 facilities measured concurrently. Aircraft and study on-site estimates were compared at six facilities measured during the field campaign, from 1-22 days apart. Each comparison was performed using two robust statistical methods that accounted for errors in both measurement techniques: Bland-Altman difference plots, and variance-weighted least-squares regression.

In the third study, described in Chapter 4, regional methane emission rate estimates predicted by aircraft mass balance flights were compared to a fully detailed, spatially and temporally resolved, bottom-up model that included methane emissions from all known sources within the region. Gridded bottom-up results were propagated downwind using Gaussian dispersion to develop simulated transects that showed similar features to longitudinal emission rate profiles measured by aircraft. Sensitivity studies were performed to understand potential contributions to the difference in means between top-down and bottom-up methods. Extensive activity data was provided by regional operators ("study partners") from all natural gas sectors, which enabled emission rate simulations at sub-hourly time scales.

CHAPTER 2

ESTIMATION OF METHANE EMISSIONS FROM THE U.S. NATURAL GAS GATHERING AND PROCESSING SECTOR¹

2.1 INTRODUCTION

A Monte Carlo model was developed that combined recent field measurements^{14,16} with facility counts obtained from state and national databases to estimate total CH₄ emissions from the U.S. natural gas gathering and processing sector. Downwind tracer flux measurements¹⁶ were made at 114 gathering stations and 16 processing plants over a 20-week field campaign¹⁴ between October 2013 and April 2014. Measurements were made in 13 U.S. states at facilities selected randomly from 738 gathering and processing facilities operated by four study partner companies. A detailed description of the downwind tracer flux measurements and accompanying data analysis are provided in Roscioli et al.¹⁶ A detailed description of the field campaign and the influence of various factors (site-type, throughput, operator presence, etc.) on measured CH₄ facility-level emission rates (FLERs) are described in Mitchell et al.¹⁴

In this study, Monte Carlo methods were used to assign CH₄ FLER values measured during the field campaign to the larger population of U.S. natural gas gathering and processing facilities, using non-parametric re-sampling, as in Ross²⁵. One model was used to estimate the CH₄ emission rate from gathering facilities, and one model was used to estimate the CH₄ emission rate from processing plants. Facility lists were compiled from various public and

¹This chapter is based on material published in *Environmental Science and Technology*: Marchese, A. J.; Vaughn, T. L.; Zimmerle, D. J.; Martinez, D. M.; Williams, L. L.; Robinson, A. L.; Mitchell, A. L.; Subramanian, R.; Tkacik, D. S.; Roscioli, J. R.; et al. Methane Emissions from United States Natural Gas Gathering and Processing. *Environ. Sci. Technol.* **2015**, *49* (17), 10718–10727, DOI: 10.1021/acs.est.5b02275. My contributions included: observing tracer measurements during the field campaign and logging field data, compiling data from state air permit databases to build input facility lists, developing the Monte Carlo simulation software, running the Monte Carlo model simulations and compiling output data sets. Consequently, this chapter includes an expanded description of the Monte Carlo simulation and preparation of input data. This chapter contains reformatted figures and tabular data that may be based on different model runs than the published article.

private inventories to build the total U.S. population. The gathering model utilized a private list of facilities provided by industry study partners, and a public list of facilities compiled from air permit databases of several gas producing states including Arkansas, Colorado, Louisiana, New Mexico, Oklahoma, Pennsylvania, Texas, and Wyoming. The processing model utilized a list of processing plants compiled from the EPA, the EIA²⁶, and the Oil and Gas Journal²⁷.

The output of each model is a cumulative distribution function (CDF) representing 50,000 possible annual CH₄ mass emissions in gigagrams (1 Gg=10⁹ gram) for the population of facilities in the input facility list. The general approach used in each of the models is as follows:

1. Compile a list of facilities which represents the overall population of facilities within the sector as completely and accurately as possible.
2. Use guided Monte Carlo methods to assign CH₄ FLER values measured during the field study to each facility in the list.
3. Sum CH₄ FLER from each facility in the list, annualize, and store the value.
4. Perturb the input facility list based on evaluated uncertainties, and repeat steps 2 and 3 for 50,000 iterations.

This approach applies to both the gathering and the processing model, but the implementation of each of the models differs, as shown in Figure 2.1 and discussed in the following sections.

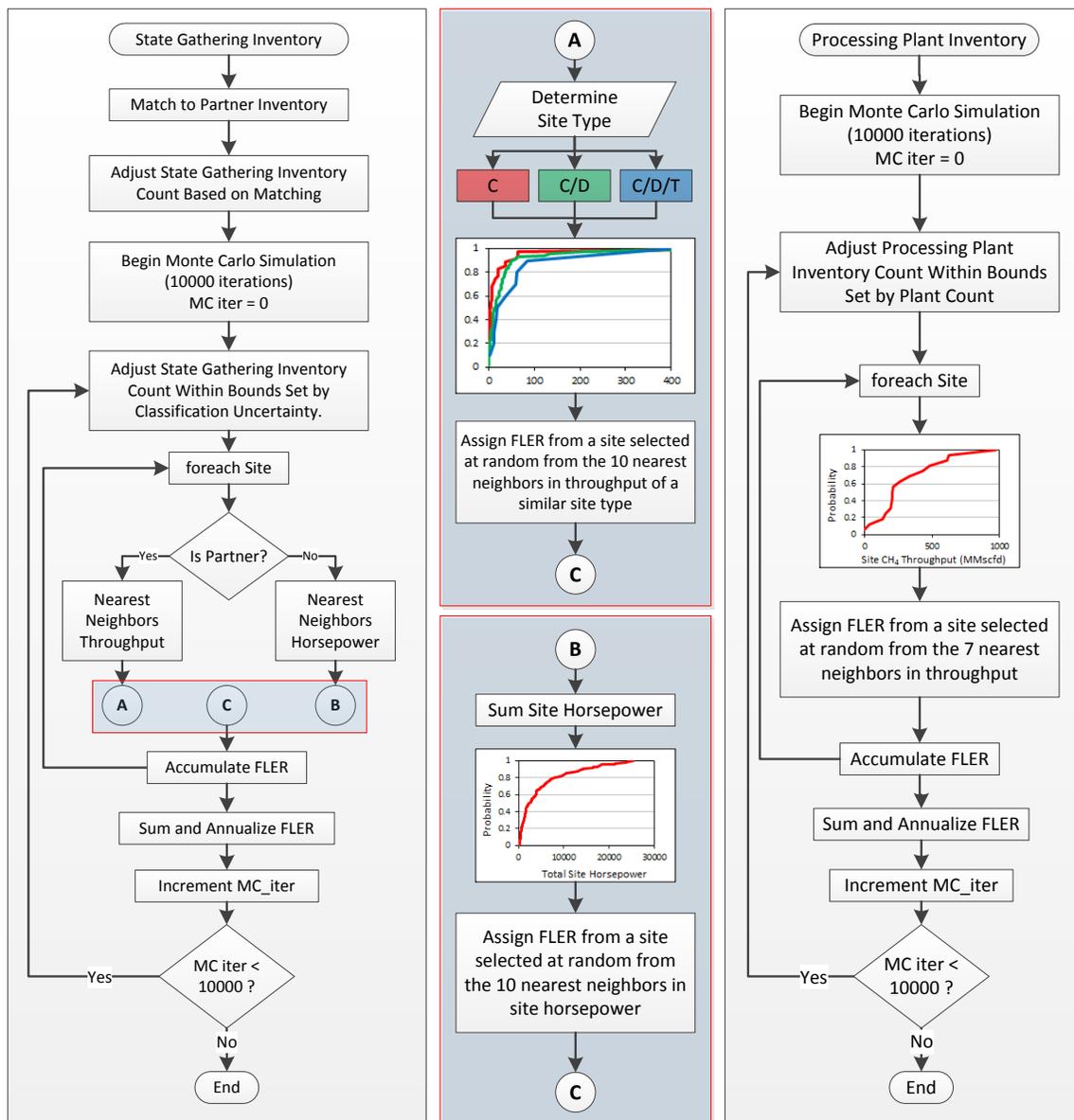


Figure 2.1: Flow chart outlining the Monte Carlo model used to predict CH₄ emissions from the gathering and processing sector. 10,000 iterations are illustrated, but 50,000 iterations were performed to develop reported results.

2.2 GATHERING MODEL

Three datasets were used as primary inputs to the gathering model to develop a CH₄ emissions estimate for the U.S. natural gas gathering sector. Data were compiled from: state air permit databases (“state data or state list”); the field study (“field data or field list”); and inventories provided by industry study partners that participated in the study (“partner data or partner list”). Information extracted from the state data (N=2,513) typically included the facility name, owner, operator, location information (address, latitude and longitude, township, section and range, UTM, etc.), and the total on-site compressor horsepower. Partner data (N=587) included similar fields, but also offered more detailed information including average natural gas throughput. Field data (N=114) included location information, on-site compressor horsepower, average facility natural gas throughput, as well as measured FLER values. The gathering model categorizes gathering facilities based on the processes performed, and the types of equipment located on-site. These site-types are defined as: compression only (C); dehydration only (D); compression and dehydration (C/D); compression, dehydration and treatment (C/D/T); and dehydration and treatment (D/T).

On each iteration of the model a facility list is created which contains gathering facilities from a state inventory. Facilities in this list are assigned FLER values from measurements obtained during the field study using Monte Carlo methods in a “nearest-neighbors” scheme. This is done in one of two ways, depending on whether a facility in the state list is identified as a partner facility. If a facility is identified as a partner facility, then the average natural gas throughput of the facility is known from partner data. In this case a FLER value will be assigned using the nearest-neighbors scheme based on site-type and natural gas throughput. If a facility in a state list is not identified as a partner facility, then the natural gas throughput is not available and a FLER value is assigned using the nearest-neighbors scheme based on

on-site compressor horsepower, which is well-correlated with facility throughput as described in Section 2.2.4.

When every facility in the list has been assigned a FLER value, the FLER values are summed and annualized, completing the iteration. This process is repeated 50,000 times resulting in a CDF of possible annual CH₄ emissions for a state. The simulation is performed for each state in the input dataset. When all states have been simulated, the gathering model is complete.

2.2.1 STATE DATASETS

Air permit data were obtained from Arkansas, Colorado, Louisiana, New Mexico, Oklahoma, Pennsylvania, Texas, and Wyoming for use in the gathering facility model. These permits include all applicable emission sources, and are not limited to the natural gas gathering sector. Thus, the first step in preparing each state inventory for use in the model is to identify which facilities in the state inventory are natural gas gathering facilities. First, facilities that are clearly unrelated to the natural gas industry are removed from the list to reduce the burden of classification. Then various methods are used to categorize the remaining facilities.

In all states, prior knowledge of the natural gas transmission and storage sector was applied to eliminate transmission and storage facilities. Facilities contained in the processing plant list described in Section 2.3.1 were removed from each state inventory. Facilities from each state inventory that were present in the partner list were classified accordingly. Classification of remaining facilities was performed based on evidence compiled using any relevant information given in the state inventory, such as Standardized Industrial Classification (SIC) code, North American Industrial Classification System (NAICS) code, company name, facility name, or Google Earth imagery.

For instance, company names were used to determine if a facility was affiliated with a company involved in the natural gas gathering industry. Company websites sometimes

include maps of their gathering pipeline networks which include facility names and locations. Facility names can be used to locate the air permit for a facility, or a public notice of the permit application. Permits often describe the function of the facility in sufficient detail to infer whether it is a gathering facility. Industrial codes can be used as well, although it is important to note that NAICS and SIC codes do not specifically identify natural gas gathering facilities, and facilities identified as gathering in this study have been listed in state inventories under a number of codes. Finally, Google Earth imagery can be used, if sufficiently detailed, to identify on-site equipment typical of a gathering facility.

2.2.2 ESTIMATION OF TOTAL STATE GATHERING FACILITY COUNTS

State gathering facility lists were compared with partner lists in an attempt to identify all partner facilities contained within each state gathering list. The comparison consisted of a two step process. First, an automated script attempted to match facilities based on company name, facility name, facility county, and distance calculated from geospatial data. After this step, the results were analyzed manually to remove false positive matches and identify any potential matches not completed by the automated script. This was necessary because naming conventions are often similar. Facilities are frequently named after the nearest town or geographical landmark. Also, geospatial coordinates were often limited in accuracy, which posed issues for the automated script in areas with densely clustered facilities.

After matching efforts were completed, the number of partner facilities found in a state gathering list was compared to the number of known partner gathering facilities within that state. The ratio of identified to known partner facilities was used to estimate the total number of gathering facilities in a state, in an effort to account for the facilities not included in state databases. For example, in New Mexico 257 facilities were identified as gathering facilities. There were 67 known partner facilities in New Mexico, and 60 were identified during the matching efforts. Thus it is estimated that the total number of gathering facilities in New

Mexico is 287. Analogous estimates were made for all other states included in the model, except for Texas, as shown in Table 2.1.

Table 2.1: State gathering facility summary data table showing gathering facilities identified in state air permit databases, partner facilities known to be located in each state, partner facilities identified in each state database, and an extrapolated gathering facility count based on the number of partner facilities identified in each state, *except for TX where an alternate method was used.

State	Identified Gathering Facilities	Known Partner Gathering Facilities	Identified Partner Gathering Facilities	Extrapolated Gathering Facilities
AR	211	63	62	214
CO	169	40	35	193
LA	312	7	6	364
NM	257	67	60	287
OK	711	45	29	1103
PA	204	58	48	246
TX	351	157	31	1012*
WY	298	150	117	382

In Texas, only 31 of 157 known partner gathering facilities were identified in the state list. Several discussions with Texas Commission on Environmental Quality (TCEQ) led the author to believe that the data received from TCEQ were complete, and that relatively poorer matching in Texas must be due to either unique reporting requirements, or a partner facility population that is not representative of all gathering facilities in Texas. An alternate method was used to estimate the number of gathering facilities in Texas, based on a recent study by Lyon et al.²⁸ They identified 259 gathering facilities in the Barnett shale region, where the total natural gas production was 11.5 Tscf from 2009 to 2014. The total natural gas production for the state of Texas was 44.9 Tscf during this same period. Using the ratio of Texas production to Barnett production, and assuming a similar number of gathering facilities per unit of gas produced, the number of gathering facilities in Texas is estimated to be 1,012.

2.2.3 ESTIMATION OF CONFIDENCE INTERVALS FOR TOTAL STATE GATHERING FACILITY COUNTS

The Monte Carlo model produces estimates of CH₄ emissions by assigning FLER values measured during the field campaign to the identified gathering facilities in the state lists. The results of the model therefore depend strongly on the number of facilities in the state lists. An analysis was performed to determine a confidence interval around the number of gathering facilities identified from state databases.

The author served as the primary observer and classified all facilities from state databases. When all facilities were classified, a random sample was extracted from each state list, and secondary observers classified the facilities based on satellite imagery. The secondary observers were not aware of the classification of the primary observer. The secondary observers were provided with geospatial coordinates, company name, and total on-site compressor horsepower for each facility in their sample. The secondary experts had spent significant time studying the satellite imagery of partner facilities, which are exemplary of gathering facilities. Therefore, they were familiar with distinguishing features inherent to gathering facilities.

A sub-sample representing approximately 10% of classified facilities was drawn at random from each state list and provided to secondary observers for classification. Facilities in the sub-sample that could not be located using the geospatial coordinates provided were not included in the comparison. The facility classifications of the secondary observers were compared to those of the primary observer, and confidence intervals were computed based on differences in classification.

For example, in New Mexico 100 facilities were drawn from the 931 left in the state facility list after reduction and classification by the primary observer. Of these 100 sites, 3 were known to the secondary observer as transmission sites. The 97 remaining sites were viewed via Google Earth, and classified by the secondary observer. The primary observer identified 14 of the 97 facilities as gathering, and 83 as non-gathering. The secondary observer's

classifications agreed with the primary observer’s for 10 of the 14 facilities classified by the primary observer as gathering. This implies that 4 out of the 14 facilities classified as gathering by the primary observer could potentially be non-gathering. Thus, the number of facilities classified as gathering by the primary observer could potentially be reduced by $\frac{4}{14}$, indicating a lower confidence bound -29%. Out of the 83 sites identified as non-gathering by the primary observer, the secondary observer classified 79 as non-gathering, and 4 as gathering. Thus, the number of facilities classified as gathering by the primary observer could potentially be increased by 4 for every 14, leading to an upper confidence bound of $\frac{4}{14} = 29\%$. This sample result can then be applied to the primary observer’s estimate on the number of gathering facilities in the state of New Mexico, leading to a confidence interval on that estimate of $\pm 29\%$.

The same process was used to determine confidence intervals on gathering facility count estimates for Colorado, Louisiana, Oklahoma, Pennsylvania, and Texas. The average of the confidence intervals for these five states were applied to Arkansas and Wyoming to save time. In states where multiple secondary observers classified sub-samples of facilities, confidence intervals were computed as described, and then averaged. Facility counts and confidence intervals are shown in Table 2.2.

Table 2.2: State gathering facility counts used in the Monte Carlo model, including confidence intervals.

State	Extrapolated		Plus CI	Minus CI	Plus Count	Minus Count
	Gathering Facilities					
AR	214		0.202	0.177	43	38
CO	193		0.240	0.180	46	35
LA	364		0.170	0.340	62	124
NM	287		0.290	0.290	83	83
OK	1103		0.120	0.120	132	132
PA	246		0.090	0.030	22	7
TX	1012		0.300	0.100	304	101
WY	382		0.202	0.177	77	68

2.2.4 FACILITY NATURAL GAS THROUGHPUT VS COMPRESSOR HORSEPOWER

Data compiled during the field campaign indicate that there is a relatively strong correlation (linear regression $R^2=0.74$) between facility natural gas throughput and operating compressor horsepower, as shown in Figure 2.2. This result supports the use of compressor horsepower as a proxy for natural gas throughput at non-partner gathering facilities where natural gas throughput is not available. However, it is impossible to determine the operating

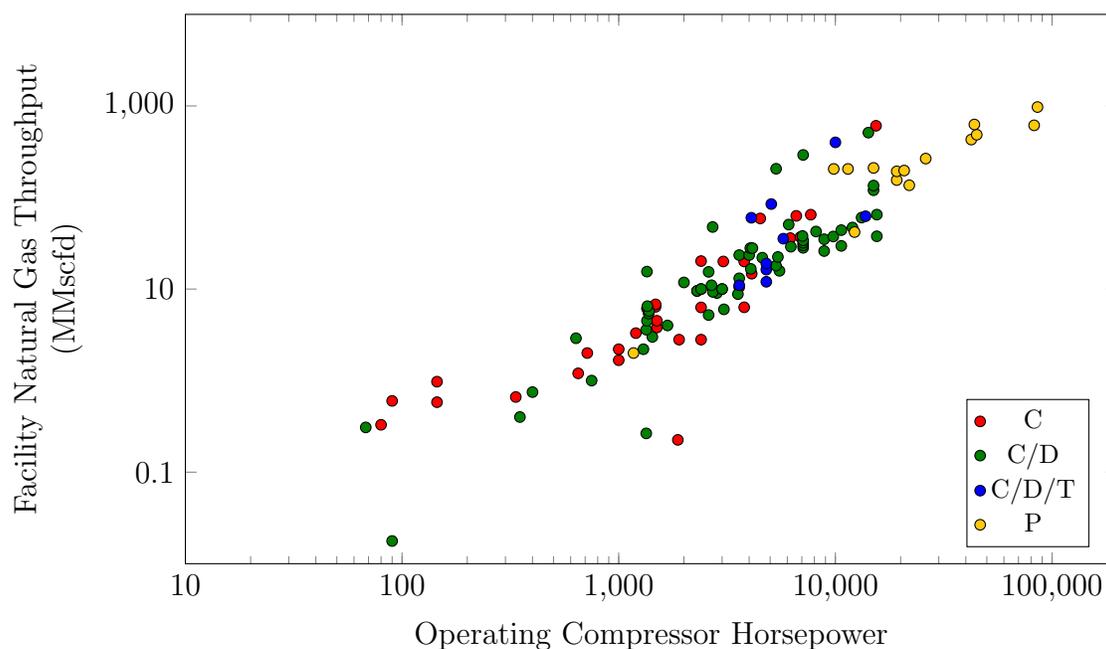


Figure 2.2: A relatively strong correlation exists (linear regression $R^2=0.74$) between actual facility natural gas throughput (MMscfd) and *operating* compressor horsepower for facilities measured during the field campaign.

horsepower of non-partner facilities in the state facility lists; only the installed compressor horsepower is available. Comparing the actual facility natural gas throughput to installed compressor horsepower at facilities visited during the field campaign also reveals a relatively strong correlation (linear regression $R^2=0.67$) as shown in Figure 2.3.

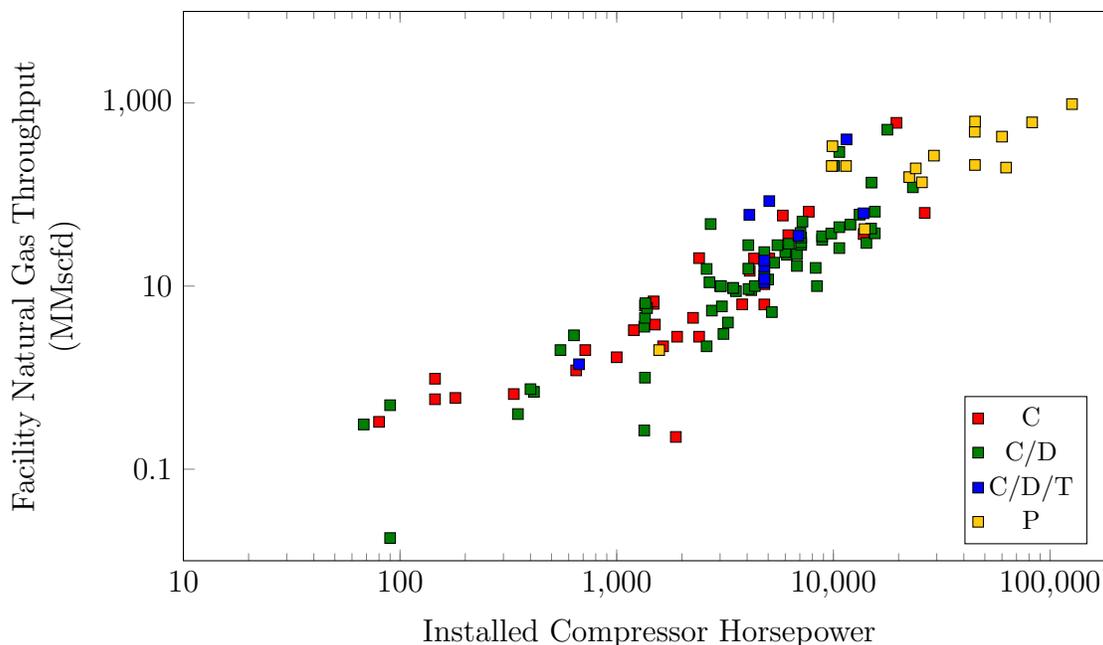


Figure 2.3: A relatively strong correlation also exists (linear regression $R^2=0.67$) between actual facility natural gas throughput (MMscfd) and *installed* compressor horsepower for facilities measured during the field campaign. This result is important as the Monte Carlo model depends on installed compressor horsepower derived from air permit data.

2.2.5 FACILITY-LEVEL EMISSION RATES VS NATURAL GAS THROUGHPUT

Field campaign data indicate that CH_4 FLER increases with increasing facility natural gas throughput, as shown in Figure 2.4. Although FLER generally increases with increasing facility throughput, there is a large variation in the FLER value measured for a given facility throughput range. This suggests that CH_4 emissions are not only affected by throughput, but are also a consequence of other factors which are independent of throughput. Mitchell et al.¹⁴ discuss both of these possibilities, noting that many factors influence CH_4 emissions, but approximately $\frac{1}{3}$ of the variance in CH_4 FLER values measured in this study can be explained by a linear regression with throughput.

All natural gas gathering facilities have similar types of equipment on-site, with similar potential for fugitive emissions. Piping, joints, flanges, valve packing, compressor seals and rod packing, liquid level controllers, pressure regulators, and other pneumatic devices all provide potential sources of fugitive emissions that are relatively steady in time. Un-burned CH_4

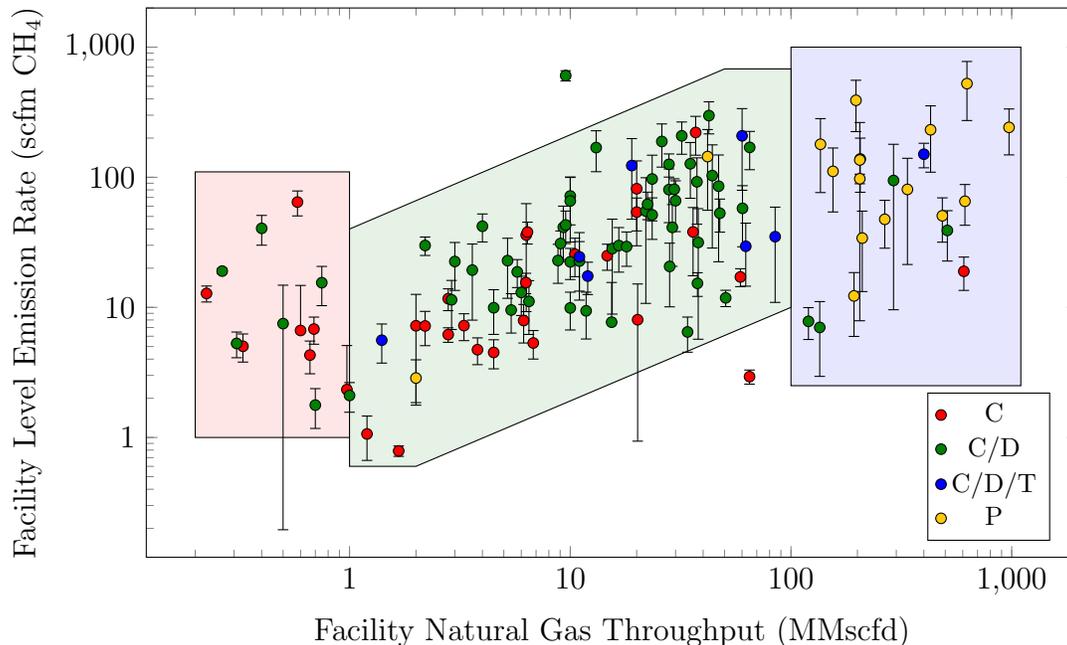


Figure 2.4: Facility-level emission rates (FLER) of CH_4 (scfm) vs facility natural gas throughput (MMscfd). Colored boxes highlight the differences in emissions for small, medium, and large facilities.

entrained in compressor engine exhaust, “combustion slip”, provides an additional source of relatively steady CH_4 emissions. Facilities with higher natural gas throughput are typically larger facilities that provide a greater number of these potential sources, leading to a rough scaling of FLER with facility natural gas throughput.

While steady CH_4 emissions are relatively easy to identify and quantify, the episodic emission sources that lead to large variation in FLER for a given facility throughput range are much more difficult to explain. The frequency, magnitude and duration of these types of emissions are unknown, and are likely influenced by a variety of factors that are difficult to quantify. Frequency of operator presence at a facility, the use of voluntary or regulated leak detection and repair (LDAR) programs, preventative maintenance practices, equipment malfunctions and failures, weather, etc., are all likely contributors to the variation in FLER within a given throughput range.

For example, Mitchell et al¹⁴ found that facilities with turbine powered compressors had lower FLER than facilities with reciprocating engine-driven compressors (75% lower on av-

erage). A portion of this reduction in emissions can be attributed to the fact that turbines generally emit less unburned fuel from the exhaust stack than reciprocating engines²⁹. However, it is likely that other factors also influence this observation. Anecdotal observations during the field campaign indicate that facilities with pressure ratio and throughput requirements suitable for the installation of a turbine-driven centrifugal compressor tend to have fewer large units, rather than several smaller reciprocating compressors. The associated reduction in piping complexity and equipment may reduce opportunities for fugitives. High-throughput facilities with turbine-driven centrifugal compressors often have staff on-site who may notice issues and flag them for repair. Issues at remote, unmanned facilities may go unnoticed for some time prior to identification and repair. None of these factors are directly quantifiable in an objective manner. The large variation in FLER vs facility natural gas throughput is addressed via the nearest-neighbors model, which randomly selects values from the “neighborhood” of facilities with similar throughput, as shown in Figure 2.6. This technique maintains the correlation of FLER with facility throughput and attempts to incorporate the influence of stochastic and difficult to quantify factors on FLER. By randomly applying measurements of similar facilities made under different circumstances, plausible FLER estimates can be made for a large number of facilities based on a smaller number of FLER measurements.

Normalizing FLER by facility throughput shows a negative correlation, as illustrated in Figure 2.5. Throughput normalized FLER (tnFLER), represents the percentage of CH₄ emitted by a facility, divided by the CH₄ throughput at the facility on the day of sampling. Gathering facilities had tnFLER greater than 10% in 4 cases, 85 gathering facilities had tnFLER lower than 0.01%. Facilities with higher tnFLER were generally smaller throughput facilities, where typical or even modest emissions were a significant portion of the facility throughput. Facilities with lower tnFLER were generally larger facilities that had greater overall emissions, but much greater throughput that overshadowed the increase in absolute emissions. Additionally, processing plants generally had lower tnFLER than gathering fa-

cilities, which may be attributable to mandatory LDAR programs and increased operator presence. As noted in Mitchell et al.¹⁴ 19 of the 25 facilities with the lowest tnFLER were staffed by operator(s) full-time (8-24 hours/day). All of the processing plants visiting during the field study were staffed full-time, while only 14% of gathering facilities were staffed full-time.

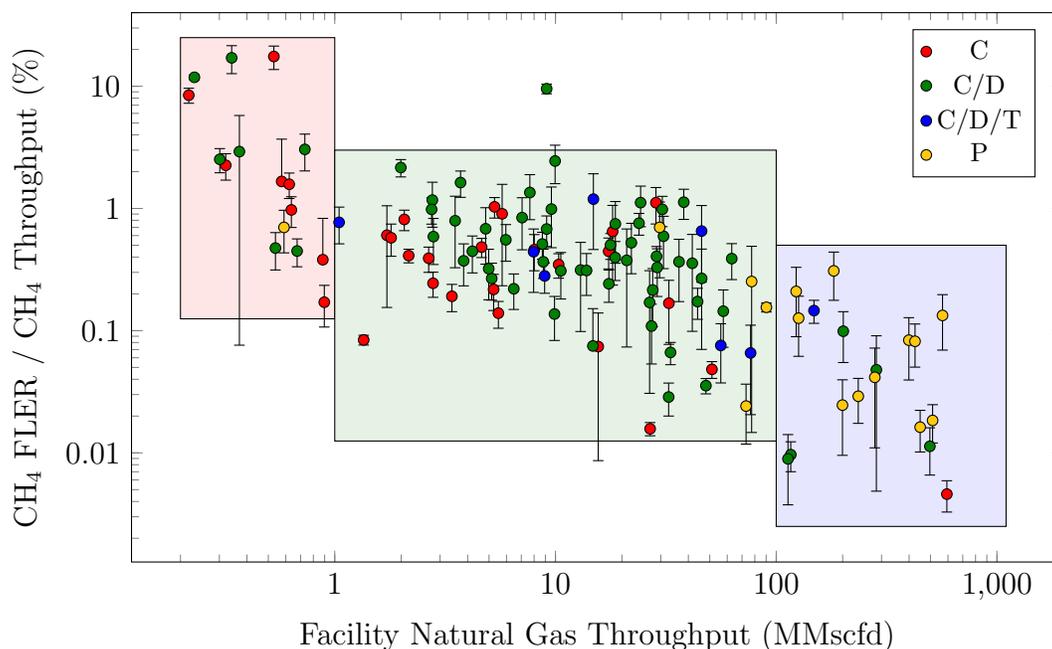


Figure 2.5: Throughput normalized facility-level emission rates (tnFLER) of CH_4 (%) vs facility natural gas throughput (MMscfd). Colored boxes highlight the differences in emissions for small, medium, and large facilities.

2.2.6 GATHERING MODEL WALK-THROUGH

In addition to facility count, the results of the Monte Carlo model are also dependent on the total on-site compressor horsepower at non-partner facilities. Typically, a state list provides a description of each compressor engine at a facility, which includes the make, model, and horsepower rating. In order to improve the accuracy of the total on-site compressor horsepower at a facility, each compressor engine listed in a state list is checked against a database as it is imported. If the make and model are not found in the database, an error is recorded. If the make and model of the compressor engine are found in the database, then

it is added to a “valid compressor list”. The valid compressor list keeps a running tally of all valid compressor engines imported, the provided horsepower of each compressor engine imported, and a running average of the provided horsepower for all compressor engines of the same make and model. Each time a new valid compressor is imported, the horsepower rating is compared to the running average horsepower of all compressor engines of the same make and model, and an error is recorded if the difference exceeds a pre-defined threshold. If a new valid compressor is imported that does not include a horsepower rating, then the running average “best rated horsepower” for that particular make and model is assigned to it. The best rated horsepower value is continuously updated as more valid compressors of the same make and model are added to the compressor list. Examination of recorded errors allows the user to correct incomplete or erroneous make, model, and horsepower ratings. Once all compressor engines at a facility have been imported, the horsepower rating of each is summed to provide a total on-site compressor horsepower value for the facility.

The Monte Carlo model estimates methane emissions from a state by assigning FLER values measured during the field campaign to each facility in a state list. On each iteration of the Monte Carlo model, a temporary state list is created by varying the facility count from the original state list within the bounds of the established confidence interval for the state. For example, in New Mexico 257 gathering facilities were identified, which extrapolated to 287 based on the number of partner facilities matched to facilities within the list. The confidence interval indicated that the facility count in New Mexico should fall between 204 and 370. Therefore, each model iteration begins by randomly choosing a site count between 204 and 370 which becomes the facility count for that iteration. A temporary state list is then created with this number of facilities. If the facility count is below 257, facilities are chosen at random from the original state list, without replacement, until the appropriate number of facilities are chosen. If the facility count is the same as the original list, the original list is simply copied to the temporary list. If the facility count is greater than 257, it is divided by the number of facilities in the original list, and rounded up to the next whole

integer. The original list is then copied this integer number of times, and the appropriate number of facilities are drawn at random, without replacement, from this expanded list. For example, if the randomly generated temporary facility count was 321, then the original state list would be copied twice, and 321 facilities would be drawn at random without replacement from a list of 514 facilities.

Once a temporary state list is generated for a given iteration of the model, one of two sub-models is used to assign a CH₄ FLER value to each facility in the list, based on whether it has been identified as a partner facility. If the facility is identified as a partner facility, the facility natural gas throughput and site-type are known. In this case a CH₄ FLER value will be chosen at random from the set of values associated with facilities of the same type, from the 10 nearest neighbors in throughput. For example, if a facility in the temporary state list is identified as a partner facility of type C/D, a FLER value will be drawn at random from a sub-sample of 10 measured C/D facilities nearest in throughput.

The second sub-model is used when the facility in the state list is identified as a non-partner facility. In this case no distinction is made between site-type, and a CH₄ FLER value is assigned at random from one the 10 nearest-neighbors based on on-site compressor engine horsepower. The total on-site compressor engine horsepower of the non-partner facility is compared to all gathering facilities measured during the field campaign. The 10 measured gathering facilities nearest in total on-site compressor engine horsepower are identified, and one is selected at random. The CH₄ FLER value associated with the selected facility is assigned to the non-partner facility in the temporary state list. This process is illustrated in Figure 2.6.

The appropriate nearest neighbor sub-model is called until each facility in the temporary state gathering list has been assigned a CH₄ FLER. Then, all CH₄ FLER values in the temporary state list are summed and annualized to a CH₄ mass emission in Gg, completing the iteration. This CH₄ mass emission value represents one possible outcome of CH₄ emissions from the gathering facilities of a given state for a year. This process is repeated 50,000 times

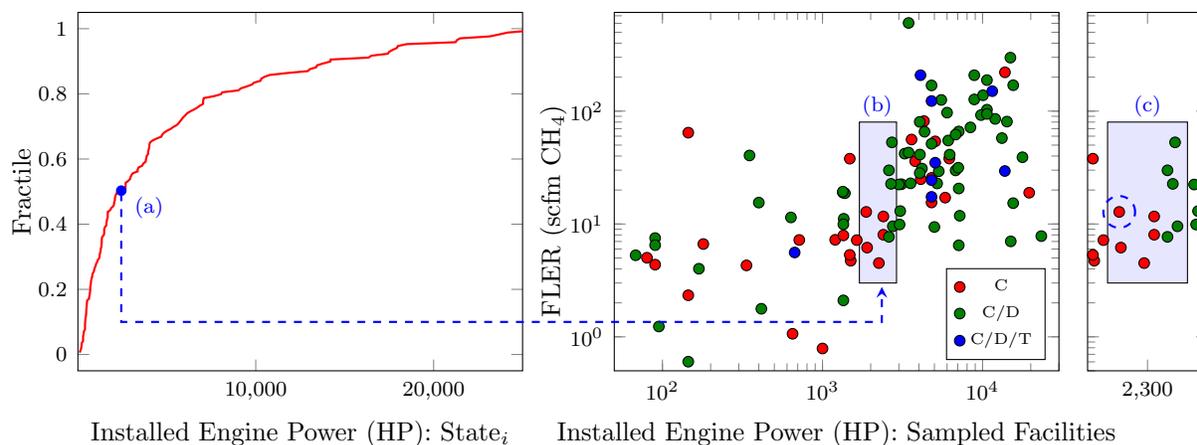


Figure 2.6: Nearest-neighbors approach to drawing a CH₄ FLER value for a non-partner facility. (a) The installed compressor engine horsepower at a gathering facility in a state list is compared to the installed compressor engine horsepower at facilities measured in the field campaign. (b) The ten “nearest-neighbors” are identified, (c) One of the 10 nearest-neighbors is selected at random, and the CH₄ FLER value associated with the selected facility is assigned to the non-partner facility in the state list.

to produce a CDF of possible CH₄ emission values for a state for a year. In this way, a central estimate is given by the 25,000th value in the CDF, and a 95% confidence interval about the central estimate is given by the 1,250th, and 48,750th values in the CDF. This data is tabulated in Table 2.3

2.3 PROCESSING MODEL

For the processing model, a list of processing plants was assembled from several data sources including the EIA, The Oil and Gas Journal, the EPA, study partners, and the state lists used for developing the gathering model. The processing model utilized Monte Carlo methods in a nearest-neighbors scheme based on average daily plant throughput to assign CH₄ FLER values measured at both processing plants, and large treatment facilities (i.e. C/D/T facilities) to the larger population of U.S. processing plants.

The processing model is run similarly to the gathering model. At the beginning of each iteration a list of processing plants is created. FLER values from the field study are assigned to each processing plant in the list using Monte Carlo methods in a nearest-neighbors scheme.

Once every plant in the list has been assigned a FLER value, the FLER values are summed and annualized, completing the iteration. This process is then repeated 50,000 times resulting in a CDF of possible annual CH₄ emissions from processing plants.

The present study defines natural gas processing plants as those engaged in the extraction of natural gas liquids from field gas in accordance with the definition provided in 40 CFR 60.631:

Natural gas processing plant (gas plant) means any processing site engaged in the extraction of natural gas liquids from field gas, fractionation of mixed natural gas liquids to natural gas products, or both.

Plants dedicated exclusively to the fractionation of natural gas liquids were considered to be outside of the natural gas supply chain and were not included in this study. However, some processing plants measured during the field study had both extraction and fractionation systems on-site, and measured FLER would capture emissions from both of these sources.

2.3.1 PROCESSING PLANT DATASET

While the list of gathering facilities had to be assembled from air permit data from individual states, more complete lists of processing plants are readily available. The EIA performs a survey of processing plants every 3 years to track U.S. natural gas processing capacity for emergency management purposes, using form EIA-757. EIA-757 data is publicly available and includes both plant name and general location, along with plant capacity and average daily plant throughput. The Oil and Gas Journal also performs a survey of U.S. natural gas processing plants. This data does not appear to provide plant capacities or throughput, but was used for activity data in the 2014 EPA GHGI. Processing plants also report to the EPA greenhouse gas reporting program (GHGRP) which is publicly available, and some processing plants were included in the state air permit data used to develop the gathering model. Industry study partners provided data for processing plants they owned or operated.

The processing plant list from the Oil and Gas Journal (N=606) was used as a starting point to assemble the input facility list for the processing plant model. This list did not provide facility natural gas throughput, and some facilities listed were found to be inconsistent with the 40 CFR 60.631 processing plant definition. These facilities were compared to the partner processing plant list (N=28), and the EIA-757 list to obtain facility natural gas throughput. If a plant had natural gas liquid (NGL) storage capacity, NGL separation equipment onsite, or was listed in EPA GHGRP data under NAICS code 211112 it was considered a processing plant as defined in this study. Of the 578 unique processing plants that were identified using this procedure, natural gas throughput was available for 512 plants. These 512 plants formed the “plant list” for the processing plant Monte Carlo model. In 2012 these 512 plants processed 16.55 Tscf of natural gas, which represents 94.7% of the 17.54 Tscf of natural gas processed in the U.S. in 2012³⁰.

The distribution of throughput values for processing plants measured in the field campaign is somewhat different than those in the plant list assembled for the processing plant model, as shown in Figure 2.7. Greater than half of the facilities in the plant list have daily average throughput values less than 30 MMscfd, while only one of the measured processing plants fell below this threshold. The number of measured processing plants is also relatively small. In order to improve both of these shortcomings, the FLER values from measured C/D/T facilities were included in the field data for the processing model. C/D/T facilities tend to be among the larger gathering facilities, and are similar to small processing plants in terms of gas throughput and on-site equipment. Additionally, although this study defined processing plants as facilities involved in the extraction of natural gas liquids consistent with 40 CFR 60.631, C/D/T facilities would be considered processing plants under EPA Subpart W reporting requirements, due to the presence of amine treatment.

The emissions predicted by the processing model were found to vary with the number of nearest-neighbors. The sensitivity of the processing model to the number of nearest-neighbors was examined for values between 2 and 12, as shown in Figure 2.8. Blue symbols

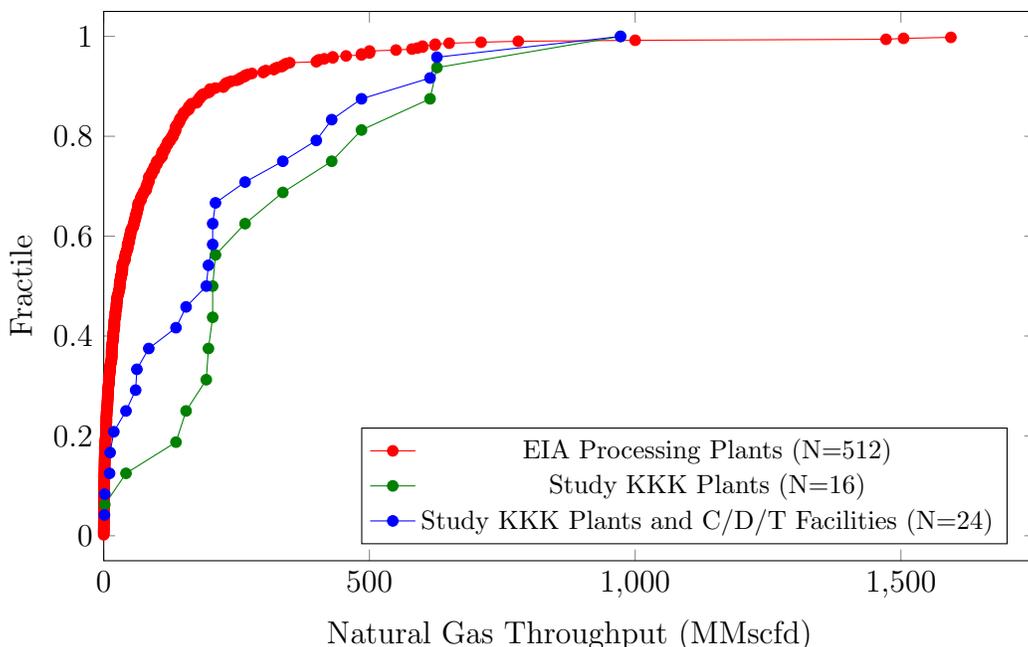


Figure 2.7: Cumulative distribution functions (CDFs) showing average daily natural gas throughput for (N=512) processing plants defined by 40 CFR 60.631, processing plants measured during the filed campaign (N=16), and combined processing plants and C/D/T facilities measured during the field campaign (N=24).

represent the median CH_4 emission predicted by the Monte Carlo model, and error bars represent the 95% confidence interval on the model prediction. The shaded blue region shows the mean (solid line), and standard deviation (dashed lines) of the predicted result for nearest neighbors values between 2 and 12. A nearest-neighbor value of 7 was chosen for the processing simulation because it closely approximates both the mean and median of tested values. The model was also tested using a nearest-neighbor value of 24, which is the same as not using a nearest neighbors model, and not taking into account the variation in emissions with throughput. In this scenario, processing plant emissions are increased to a point where smaller processing plants would have unrealistically high tnFLER values. For example, 20% of the modeled processing plants have daily throughput less than 4 MMscfd, and assigning the average FLER value from processing plants measured by Mitchell et al.¹⁴ (181 kg/h) to these facilities would result in tnFLER values of 5% and greater.

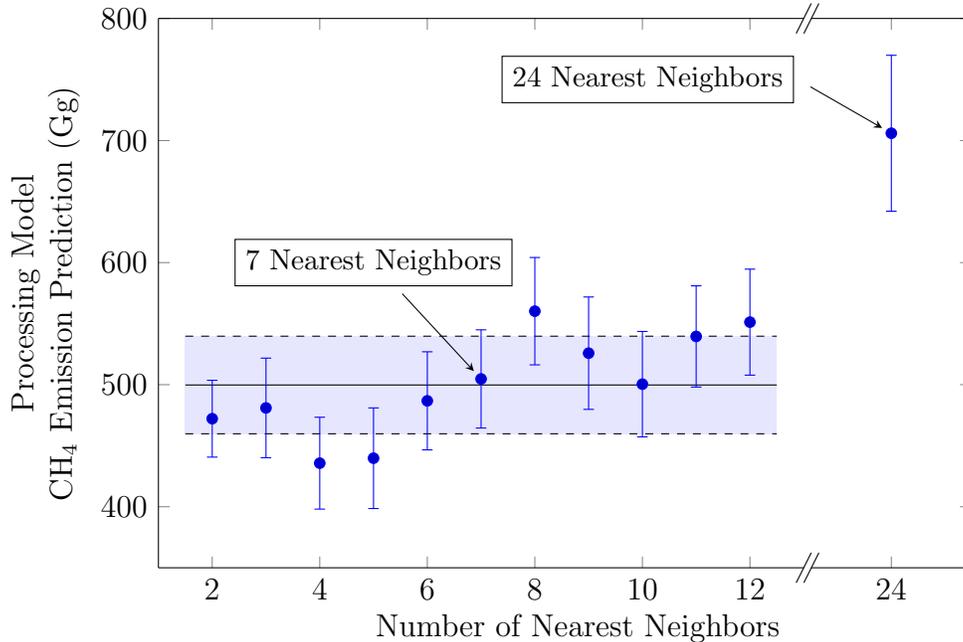


Figure 2.8: Effect of number of nearest neighbors on predicted CH₄ emissions from the processing plant model. 7 nearest-neighbors were used in the final model.

2.3.2 PROCESSING PLANT MODEL WALK-THROUGH

Like the gathering model, the CH₄ emissions predicted by the processing model are also dependent on the processing plant count used in the model. The processing plant Monte Carlo model uses a facility count of 592 ± 14 , which represents the average of the upper and lower bounds provided by the facility count from the Oil and Gas Journal (N=606), and the number of unique 40 CFR 60.631 processing plants (N=578) identified in this study. On each iteration of the Monte Carlo model, the facility count is varied by $\pm 2.4\%$ by choosing a value at random from within this range.

At the beginning of each iteration, a temporary plant list of N=1024 facilities is created by duplicating the list of 40 CFR 60.631 processing plants with natural gas throughput available (N=512). A facility count is generated, and the appropriate number of facilities are drawn at random without replacement from the temporary list, forming the input plant list for the processing plant model. The natural gas throughput of each facility in the input plant list is compared to the natural gas throughput of the 24 P and C/D/T facilities measured in the

field campaign. The 7 nearest neighbors in natural gas throughput are identified, and one is selected at random. The CH₄ FLER value associated with the selected facility is assigned to the processing plant from the input plant list. Each processing plant in the input plant list is assigned a CH₄ FLER value in this way. The assigned CH₄ FLER values are then summed and annualized, giving one possible value of annual CH₄ emissions from processing plants. This process is repeated 50,000 times, resulting in a CDF of possible emission values. In this way, a central estimate is given by the 25,000th value in the CDF, and a 95% confidence interval about the central estimate is given by the 1,250th, and 48,750th values in the CDF as shown in Figure 2.10.

2.4 MONTE CARLO MODEL RESULTS

Monte Carlo simulation results for gathering stations in each of the eight modeled states are shown in Figure 2.9 and Table 2.3. Results for individual states are reported as the median probability (50th%) of simulation results for each state. The upper and lower uncertainty reported is the difference between the median simulation result and the 2.5th% and 97.5th% results, respectively. This range represents a 95% confidence interval about the median value. Sums provided in Table 2.3 are reported similarly, but for sums calculated after each iteration. For this reason, the “Sum of States” total differs from the total of individual state results shown in Table 2.3. A CDF of simulation results for each individual state is shown in Figure 2.9. Since each state is simulated from the same 114 FLER measurements made during the field campaign, total emissions and associated uncertainty are directly related to site count and site count uncertainty. The slope of each CDF indicates the uncertainty in site count. For example, Colorado and Pennsylvania have similar total emissions, but Colorado has a larger 95% confidence interval due to greater uncertainty in site count. Uncertainty is also influenced to a lesser extent by the uncertainty in individual FLER measurements.

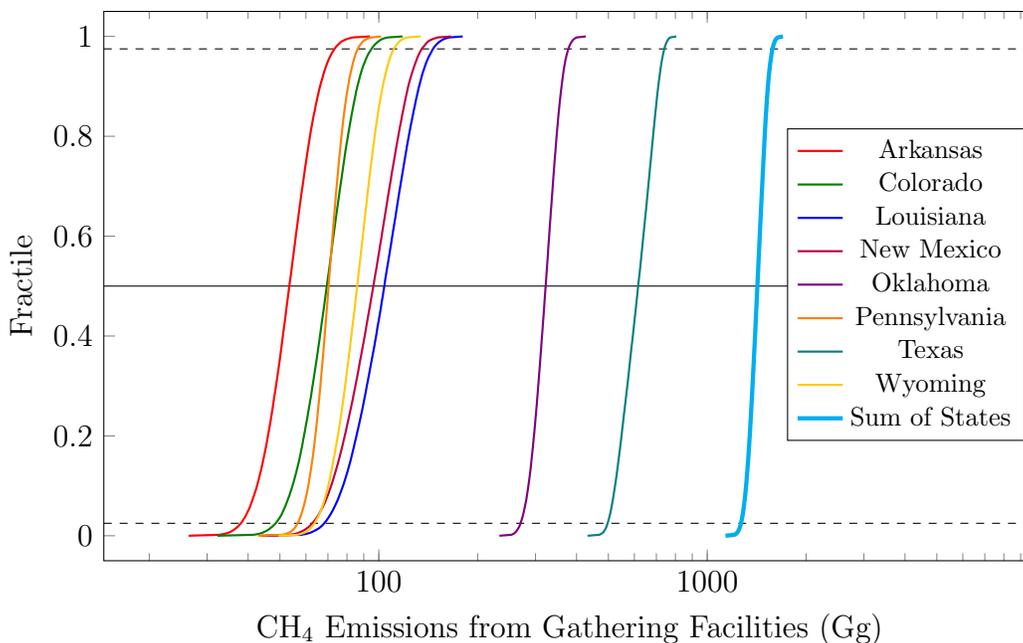


Figure 2.9: Cumulative distribution functions of predicted annual emissions of CH_4 (Gg) from gathering facilities in the states listed, for 50,000 iterations of the Monte Carlo model. The median emission value predicted by the model is shown by the intersection of each CDF with the solid line (50th%). The lower and upper confidence bounds are shown by the intersection of each CDF with the dashed lines (2.5th% and 97.5th% respectively).

Total CH_4 gathered for each state was calculated by subtracting lease fuel from total marketed production based on EIA data³¹ and assuming an average CH_4 content of 90% in gathered gas, as described in Marchese et al.¹⁵ A modeled CH_4 loss rate was then calculated for each state by dividing the total CH_4 emissions by the total gas gathered. Oklahoma and Texas have the largest absolute emissions of simulated states; however, Oklahoma is the only state with a modeled CH_4 loss rate greater than 0.5% due to a greater proportion of lower throughput facilities than other states. For example, 96% of facilities in Oklahoma have installed compressor horsepower less than 7,500 hp, while in Arkansas only 70% of facilities have installed compressor horsepower less than 7,500 hp. Simulated CH_4 emissions from the eight modeled states totaled 1,421 Gg (1,287–1,558 Gg 95% CI), which represents an eight-state modeled CH_4 loss rate of 0.4%. These eight states accounted for 83.5% of the U.S. total (351,310 Gg of 420,906 Gg) CH_4 gathered in 2012³¹. If the facilities that gather the remaining 16.5% of CH_4 experience the same loss rate, 1,699 Gg (1,539–1,863 Gg

95% CI) of CH₄ emissions are predicted from U.S. gathering facilities. The total number of U.S. gathering facilities was also estimated at 4,554 (+434/-287) by assuming that facility count and relative uncertainty scale directly with the quantity of CH₄ gathered. If these assumptions hold, the average U.S. gathering facility is predicted to emit 43 kg/h (+6/-5 kg/h) CH₄. Uncertainties include the 95% confidence interval about the median eight-state modeled emissions, and the uncertainty in state site count developed in Section 2.2.3. The Monte Carlo model result for U.S. processing plants predicts 505 Gg (+43/-46) Gg, as shown in Figure 2.10. This estimate is inclusive of all known U.S. processing plants, and does not require additional scaling.

Table 2.3: Monte Carlo simulation results for modeled states, including confidence intervals. Modeled CH₄ loss rate calculated from simulated emissions and CH₄ gathered by state. U.S. total estimates described in text. Site count shown for reference.

State	Simulated CH ₄ Emissions (Gg)	95% CI Range (Gg)	CH ₄ Gathered (Gg)	Modeled CH ₄ Loss Rate (%)	Site Count
Arkansas	53	40–70	19 723	0.27	214
Colorado	69	51–91	28 261	0.25	193
Louisiana	104	72–140	50 207	0.21	364
New Mexico	96	67–130	20 215	0.47	287
Oklahoma	322	277–369	34 263	0.94	1103
Pennsylvania	70	59–84	37 676	0.19	246
Texas	616	509–727	126 552	0.49	1012
Wyoming	86	67–107	34 414	0.25	382
Sum of States	1421	1287–1558	351 310	0.40	3801
US Total est.	1699	1539–1863	420 906	0.40	4554

Results predicted by the Monte Carlo simulations may be biased by several factors, including: emissions from episodic sources; incomplete capture of combustion slip entrained in lofted, buoyant exhaust plumes; and uncertainty in the count of electric compressor stations. During the field campaign, ten individual tracer plumes captured episodic emissions from compressor blow-downs, start-ups, and a similar event of unknown origin. These high-emission rate plumes were excluded from FLER estimates, which are therefore biased low. Marchese et al.¹⁵ developed a separate Monte Carlo model to estimate the impact of excluded episodic sources and found that the gathering facilities may emit an additional 169 (+426%/-

96%) Gg of CH₄ annually. Roscioli et al.¹⁶ performed Gaussian dispersion modeling with Brigg’s plume rise equations and found that significant portions combustion slip emissions may not be recovered under certain scenarios. Mitchell et al.¹⁴ tested several “worst-case” scenarios and estimated that low-bias due to incomplete recovery of combustion slip is likely at 5-20% of measured facilities. The influence of electric compressors station is unclear. Emissions from these sources were not accounted for in the Monte Carlo model; therefore, the results presented here are likely biased low.

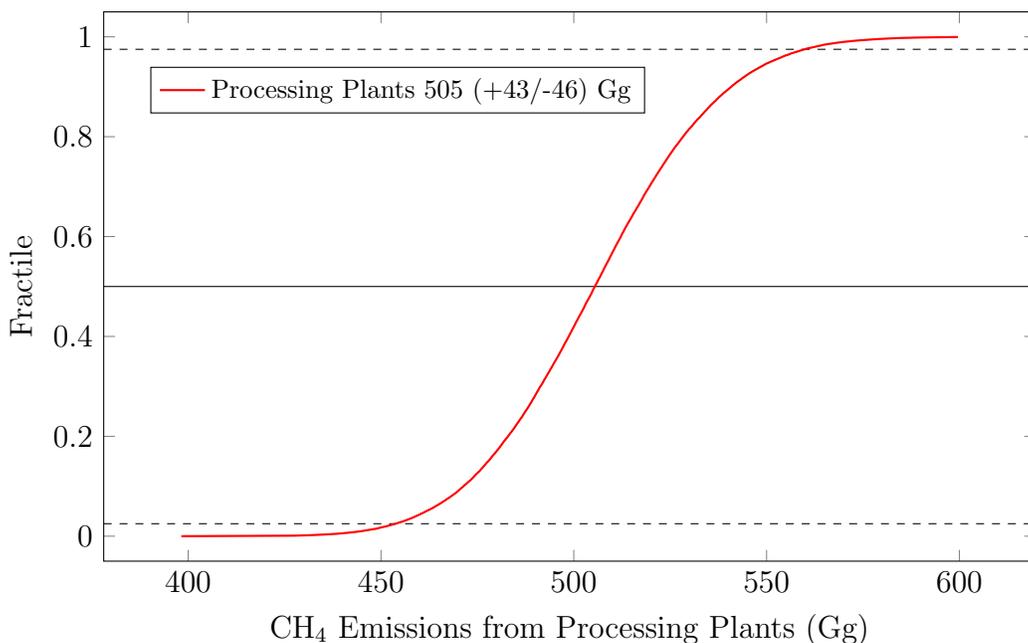


Figure 2.10: Cumulative distribution function of predicted annual emissions of CH₄ (Gg) from U.S. processing plants.

2.5 CONCLUSIONS

Marchese et al.¹⁵ provide detailed comparisons of the Monte Carlo model results to two national inventories: the EPA GHGI, and the EPA GHGRP. The GHGI³² includes natural gas processing as a distinct sector within ‘Natural Gas Systems’, allowing a direct comparison with modeled processing plant results. Gathering activities are classified within ‘Natural Gas Systems’ under ‘Field Production’, which required disaggregation of gathering activities from production activities before a comparison to modeled results could be made. Emissions from

activities unique to the gathering sector were assigned directly to gathering, and emissions from activities common to both gathering and production were apportioned based on partner data and the estimated U.S. gathering facility count developed herein. Both gathering and processing facilities report to the GHGRP under various subparts, but only if their total greenhouse gas emissions are greater than 25,000 tonne/yr of CO_{2e}. Processing plants report under Subparts C and W, while gathering facilities only report under Subpart C for stationary combustion emissions.

Modeled processing plant emissions (505 Gg (+43/-46 Gg)) were lower than net processing plant emissions estimated in the GHGI (892 Gg). The GHGI estimate includes 40 Gg of CH₄ emissions attributable to maintenance activities, a category not accounted for in the Monte Carlo model. The remaining difference is likely due to disparate activity data for compressor units. 90% of processing plant CH₄ emissions in the GHGI are due to compressor fugitives and combustion slip; the GHGI lists an average of 9.3 reciprocating compressors per plant, and a ratio of 6.2 reciprocating to centrifugal compressors. Lower reciprocating compressor counts and ratios were observed at processing plants (6.4 and 2.6) and C/D/T facilities (5.4 and 3.3) during the field campaign. This may be an indication that plants are shifting from numerous reciprocating compressors, to fewer, larger centrifugal units.

Gathering sector emissions disaggregated from the GHGI field production sector totaled 404 Gg, and included 178 Gg from gathering pipelines, a category not included in the Monte Carlo model. Subtracting pipeline emissions results in a GHGI estimate of 226 Gg from gathering facilities alone. Total U.S. gathering facility CH₄ emissions predicted by the Monte Carlo model (1,699 Gg (1,539–1,863 Gg 95% CI)) exceed this amount by a factor of seven, and the model prediction is likely biased low.

The 2013 GHGRP estimates 0.53 Gg of CH₄ emissions from 404 reporting gathering facilities, and 179 Gg from 433 reporting processing plants. Both of these estimates are much lower than emissions predicted by the Monte Carlo model due to: lower facility counts resulting from the reporting threshold, an unrealistically low combustion slip emission factor

for reciprocating compressor engines, and reporting requirements that exclude emissions from certain source categories.

Comparing results predicted by the Monte Carlo model to each of these inventories highlights several important points:

1. Emissions from the gathering sector are currently under-estimated in inventories, but represent a significant portion of CH₄ emissions from natural gas systems.
2. The GHGI may currently over-estimate emission from processing plants due to activity data that does not reflect current operating practices.
3. The GHGRP does not provide a complete assessment of CH₄ emissions due the reporting threshold, inaccurate emission factors, and categorical exclusions.

CHAPTER 3

COMPARING FACILITY-LEVEL METHANE EMISSION RATE ESTIMATES AT NATURAL GAS GATHERING AND BOOSTING STATIONS²

3.1 INTRODUCTION

Gathering systems use pipelines to collect gas from upstream wells and deliver it to gathering and boosting stations (hereafter “gathering stations”). Gathering stations include natural gas compressor equipment that boosts the pressure of the produced gas from well pressure to the required downstream pressure. Compressors are typically driven by reciprocating engines fueled by a fraction of the gas passing through the station. Gathering stations may also be equipped with a range of supporting equipment, including dehydrators to remove water, treating equipment to remove undesirable gases, fuel conditioning systems, piping and control lines, metering, and other associated support equipment. Stations discharge gas to downstream pipeline networks that feed processing plants, transmission systems, distribution systems, or other gathering stations. Figure 3.1 provides an equipment overview of a typical Fayetteville gathering station measured in this study. Emissions from gathering pipelines outside of the gathering station boundary are not considered in this study, but are addressed in a companion study performed during the same field campaign by Zimmerle et al.³³

Recent studies have used a variety of measurement methods to quantify CH₄ emission rates from natural gas systems.^{34,35} Top-down methods that rely on atmospheric CH₄ mole fraction measurements alone may have difficulty attributing emissions to distinct sources, e.g., from biogenic or thermogenic sources at the regional scale, or from individual point sources at the facility scale.³⁶ Tracer-release measurements help address the latter issue by

²A version of this chapter has been submitted as: Vaughn, T. L.; Bell, C. S.; Yacovitch, T. I.; Roscioli, J. R.; Herndon, S. C.; Conley, S.; Schwietzke, S.; Heath, G. A.; Pétron, G.; Zimmerle, D. Comparing Facility-Level Methane Emission Rate Estimates at Natural Gas Gathering and Boosting Stations. *Elem. Sci. Anth.* **2017**.

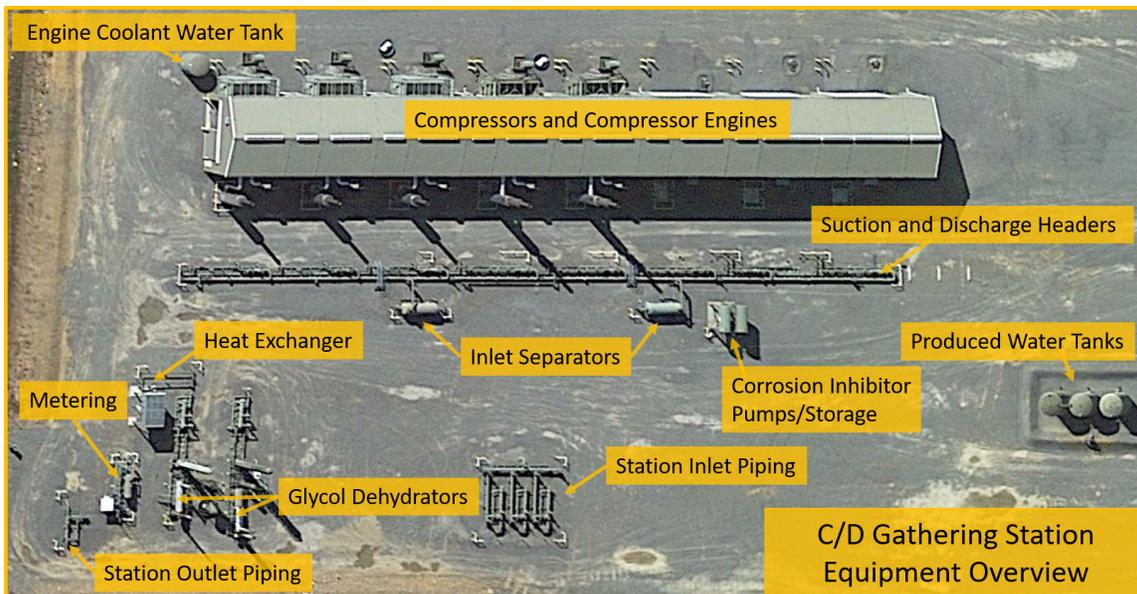


Figure 3.1: A typical gathering station in the Fayetteville shale. This station features compression and dehydration (water removal).

providing confirmation that CH_4 emissions are co-located with tracer gas release points; however, tracer measurements may not capture all emissions from a facility under certain conditions.¹⁶ A bottom-up estimate based on direct measurements of all components at a facility is possible in theory, but is challenging in practice due to the large number of potential sources, and difficulty of measuring or accurately estimating every source. Certain direct measurements may not be possible due to personnel safety or accessibility issues. A variety of factors may contribute to differences between top-down and bottom-up estimates of CH_4 emissions. Recurring themes in recent discussions and studies^{34,35} include temporal variability, unrepresentative emission factors, and skewed emission rate distributions.

The work presented here is part of a large, multifaceted field campaign designed to estimate methane emissions from all segments of the natural gas supply chain within the study area (production, gathering and boosting, transmission, and distribution systems). Multiple contemporaneous measurement methods were used to develop estimates of methane emissions at the device, facility and regional-scale to help reconcile top-down and bottom-up estimates. Here, we compare CH_4 FLER estimates developed from on-site, dual-tracer

and aircraft methods made either concurrently (same day, same time) or contemporaneously (both made during this field campaign) at gathering stations. Estimates for unmeasured sources were based on other measurements made in this study, if possible, or prior measurements of specific sources. Facilities with emission sources well outside the measurement capability of a method were excluded from comparisons.

3.2 FIELD STUDY

Measurements for this study were collected during a four-week field campaign conducted in September–October 2015 in an eastern portion of the Fayetteville shale play (the “study area” Figure 3.2). Fayetteville gas contains little hydrogen sulfide or other trace gas contaminants and few hydrocarbons heavier than ethane. Consequently, the gas is considered “sweet and dry” and requires only dehydration and acid gas removal prior to sale. There are no processing plants or storage facilities within the study area; gas discharged from gathering stations is routed directly to transmission or distribution systems. Ninety-nine (79%) of the 125 gathering stations in the study area are owned and managed by companies who provided site access and supported or participated in measurement activities. These companies (“study partners”) also provided activity data, compressor engine exhaust stack test data, and insight into their operations that was critical to accurate modeling and interpretation of facility-scale emission estimates.

The field measurement campaign was designed to maximize the number of paired measurements between methods by clustered sampling. Limited road access, coupled with consistent north and northeasterly winds, made some gathering stations in the clustered plan inaccessible for measurement by dual-tracer. Therefore, during the campaign additional facilities outside the clusters were measured to maximize the total number of measurements. As a result, site selection was randomized by the combination of wind direction and road access, and over the 4-week campaign, teams measured essentially all facilities with nearby roads suitable for dual-tracer. To test for the possibility of bias, a two-sample Kolmogorov-

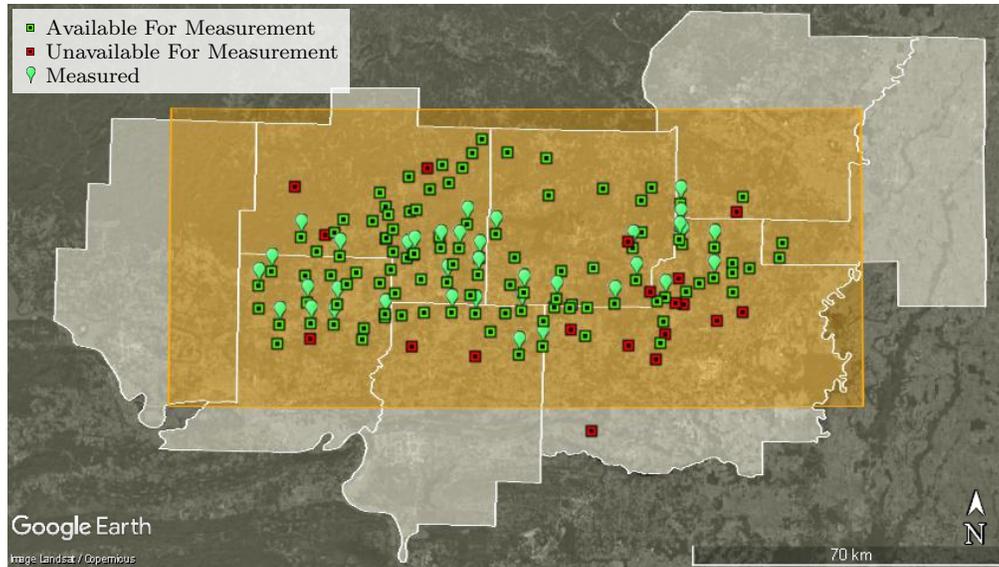


Figure 3.2: Ninety-nine out of 125 gathering stations within the study area (orange highlighted region) were available for measurement. Collectively, thirty-six stations were measured by, on-site, tracer, and aircraft teams.

Smirnov test on the size of the measured gathering stations indicates that the measured facilities represent an unbiased sample of the facilities available for measurement as shown in Figure 3.3.

During the field campaign, personnel from Colorado State University (CSU) served as study coordinators and on-site observers who coordinated measurement teams and observed measurement activities. Each evening study coordinators used wind forecasts to identify three to six potential facilities to measure the following day. This list of potential facilities was shared with study partners and measurement teams that evening or on the morning of the measurement day. During the day, the study coordinator and tracer measurement team chose measured facilities from the list of potential facilities based on the observed local wind conditions, without input from study partners.

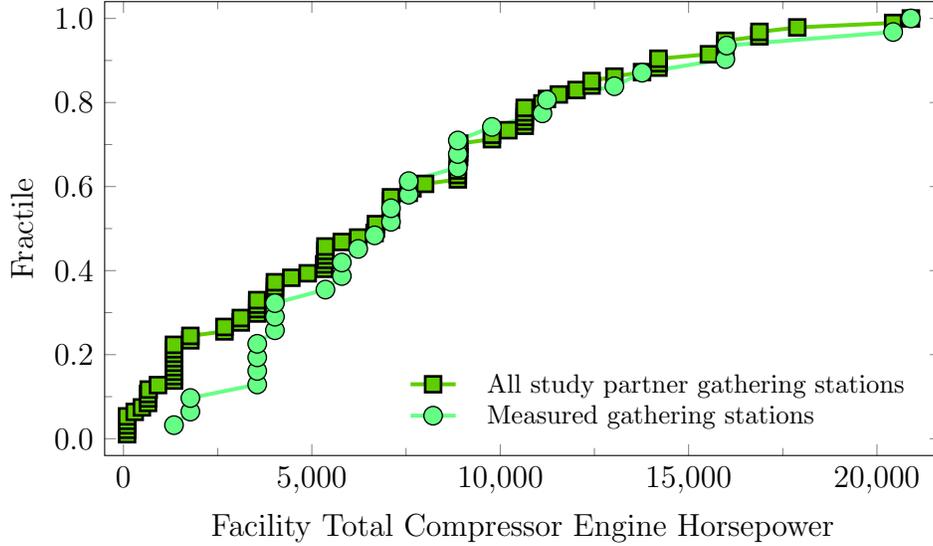


Figure 3.3: Considering the size of the stations, as estimated by installed compressor engine power, measured facilities are representative of all facilities available for measurement.

3.3 MEASUREMENTS

Independent teams using dual-tracer, aircraft spiral flights, and on-site measurement methods measured 36 unique gathering stations during the campaign. While efforts were made to maximize the number of paired measurements, scheduling constraints limited the number of facilities measured concurrently by different measurement teams. Paired measurements were not possible between tracer and on-site teams in weeks one and four, as shown in Table 3.1.

Table 3.1: Measurement team availability during the field campaign.

Team	Campaign Week			
	1	2	3	4
On-site Observer	✓	✓	✓	✓
AECOM/LDAR	✗	✓	✓	✓
Tracer Team	✓	✓	✓	✗
Aircraft Team	✓	✓	✓	✓

Teams completed measurements and consolidated results independently. On-site observers from CSU accompanied measurement teams during the field campaign, coordinated teams to obtain contemporaneous measurements, and ensured compliance with measurement

protocols. Observers also noted unmeasured emission sources, changes in facility operations, and observed measurements made by on-site teams using an optical gas imaging (OGI) camera. On-site measurements were made by AECOM Inc., or AECOM and study partner personnel “on-site or on-site team”. Tracer measurements were made by Aerodyne Research Incorporated “tracer or tracer team”, and Aircraft measurements were made by Scientific Aviation Incorporated “aircraft or aircraft team”.

Measurements were not successful at all facilities by all methods. To support valid CH₄ FLER estimate comparisons by paired methods, measurements must meet all quality control requirements imposed by the measurement team’s protocols. Additionally, facilities must remain in a uniform operating state during each measurement. The reduction from all attempted measurements to measured facilities where valid comparisons could be made is shown in Table 3.2.

Table 3.2: Counts of facilities measured during the field campaign. While measurement teams reported successful measurements at most attempted facilities, paired comparisons were not possible at all facilities, either due to lack of successful paired measurements or changes in the operating state of the facility during or between measurements.

	On-site	Tracer	Aircraft
Attempted Measurements	33	32	11
Successful Measurements	32	30	10
Possible On-site Comparisons	-	26	10
Accepted On-site Comparisons	-	21	6
Concurrent On-site Comparisons	-	14	0

In order to enable fair and appropriate method comparisons, a high level of confidence that measurements were made when facilities were in a similar emitting state is required. While there is no way that this can be proved conclusively, a best-effort attempt was made to identify happenings observed that would negate a fair comparison. This was done by noting field observations made by all measurements teams and on-site observers that might indicate a biased assessment of FLER by any given method. Additionally, operations data provided by both study partners and data partners was used to identify possible confounding factors at measured or nearby facilities. Tracer to study on-site estimate (SOE) method comparisons

were limited to facilities where an on-site observer was present during both measurements. During on-site measurements at gathering stations two on-site observers were present and equipped with an OGI camera to observe on-site measurements teams and identify any unique or unusual emissions sources or events. For aircraft to SOE method comparisons it was not possible to enforce the criterion that an on-site observer be present during both measurements due to the small sample size of paired measurements available. Facilities excluded from method comparisons are described in detail below.

Gathering Station 61

At gathering station 61, substantial portions of the facility were not covered for leak detection via OGI. Therefore, SOE is not accepted at this facility due the potential for unidentified emission sources which would have contributed to the tracer measurement, and not the SOE, preventing a fair comparison. Therefore, this facility is eliminated from the:

- Tracer to Study On-site Estimate method comparison

Gathering Station 121

The aircraft identified significant emissions from gathering station 121 during a raster flight. Gas was venting from a produced water tank, which originated from an open manual (hand-operated) dump valve on a compressor engine fuel scrubber. On-site teams were unable to measure emissions from the tank, which were above the measurement range of the high-flow sampler. The tracer team was not able to provide a tracer facility estimate (TFE) for the entire facility due to poor winds and downwind road access. However, the tracer team was able to isolate the portion of the facility where the tank was located, both with the valve open, and after it had been identified and closed. Subtracting the tracer estimate made in each operating state, and subtracting the associated uncertainties (95% confidence interval (CI)) in quadrature, leads to an estimated 606 (± 278 kg/h) of CH₄ emissions from the tank. Aircraft facility estimates were performed at this facility on three different days:

on two days measurements captured the facility in a higher emitting state 676 (± 119 kg/h), and 739 (± 107 kg/h), and on one day at a lower emitting state 276 (± 99 kg/h). If the tank emissions estimated by tracer are added to the SOE at this facility 109.6 (-8.1/+7.9 kg/h), the SOE compares well with the aircraft facility estimate (AFE) on the two days AFE captured the facility in a higher emitting state. This facility was selected for measurement by directed, and not random sampling. The on-site team was not able to measure the tank emissions; the tracer team was not able to produce a full TFE due to poor wind conditions; the aircraft captured the facility in multiple (high) emitting states, showing high variability. Aircraft measurements were not made concurrently with tracer or on-site measurements. Therefore, this facility is excluded from the:

- Tracer to Study On-site Estimate method comparison
- Aircraft to Study On-site Estimate method comparison

Gathering Station 33

The tracer team noted significant emissions from a produced water tank, which were above the range of the high-flow sampler, and which the on-site team did not attempt to measure. Study partner company operators suspected a stuck dump valve, but were unable to identify the source while measurement teams were on-site. At this facility, tracer measurements would include emissions from the tank. The on-site team was unable to measure the tank and had no way to estimate emissions with any degree of certainty. Therefore, this facility is excluded from the:

- Tracer to Study On-site Estimate method comparison

Gathering Station 111

During measurements at gathering station 111, an operating compressor was accidentally shut down, and operators experienced difficulty restarting it due to water in the fuel line. The fuel line was vented and purged, and the compressor piping was vented and purged multiple times. After several attempts the compressor engine was restarted, and normal operations resumed. On-site teams did not measure these large, non-continuous emissions, and tracer teams captured them, reporting highly variable emissions, with periods of high and unsteady concentration enhancements seen just downwind of the facility. The aircraft measured this facility 30 minutes after the compressor engine was restarted. Therefore, this facility is excluded from the:

- Tracer to Study On-site Estimate method comparison
- Aircraft to Study On-site Estimate method comparison

Gathering Station 96

It was determined from a post-campaign activity data survey that a manual liquid unloading had occurred at a well within the flight path during an aircraft spiral flight targeting a nearby gathering facility, as shown in Figure 3.4. These emissions would have contributed to the AFE, and not the TFE or SOE. Therefore, this facility is excluded from the:

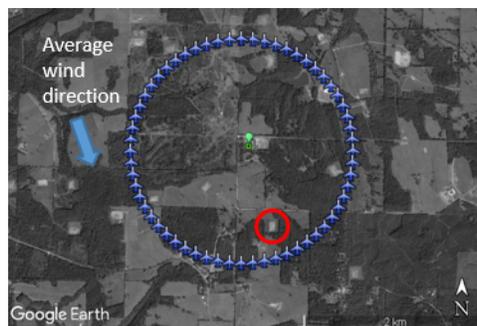


Figure 3.4: A manual liquid unloading occurred at a nearby well (red circle) during aircraft measurements targeting a gathering station at the center of the aircraft flight (green balloon).

- Aircraft to Study On-site Estimate method comparison

During post-campaign quality control, it was determined that completion work was being performed on a well immediately upwind from the aircraft flight path, as shown in figure 3.5. These emissions provided a confounding upwind source for the aircraft. Therefore, this

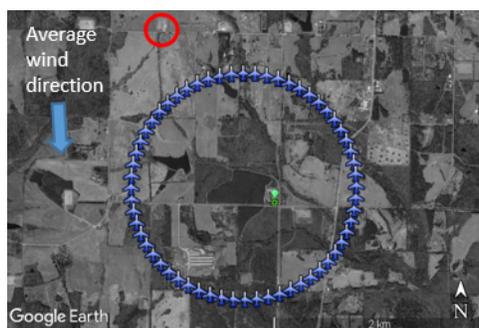


Figure 3.5: Completion work was being performed at a nearby well (red circle) that confounded aircraft measurements targeting a gathering station (green balloon).

facility is excluded from the:

- Aircraft to Study On-site Estimate method comparison

3.3.1 TRACER MEASUREMENTS

Tracer release is an established atmospheric measurement technique^{13,14,17,20} that estimates an emission rate by releasing a tracer gas at a known rate near an emission source and comparing concentrations of the target analyte (CH_4 in this case) and tracer gas downwind of the facility. Dual-tracer gasses can be released to provide an internal standard and empirical measure of uncertainty for each measured plume.¹⁶ In this study, individual FLER estimates were based on 2–14 dual-tracer plume measurements made 200–1200 m downwind of each gathering station. For a detailed discussion of the methods and results of the tracer-based FLER estimate and associated uncertainty used in this study see Yacovitch et al.³⁷ Aerodyne Research Incorporated performed all tracer measurements in this study.

3.3.2 AIRCRAFT MEASUREMENTS

Aircraft spiral flights are a new aircraft-based method for estimating emissions from individual facilities by sampling and analyzing the emission plume from the facility. The aircraft begins circling the target facility on a 900–1500 m fixed radius course, 75–175 m above ground level. The aircraft continues upward, flying loops in 100 m increments, until enhancements of the target species are no longer observed, which occurred at a height of 280–580 m in this study. Multiple loops may be flown at one level before climbing to the next. In-situ measurements of methane, ethane, wind speed, and wind direction are made during flight. The mass flux of CH₄ is calculated for loops at a given altitude using Gauss’ method. Results from each altitude are then integrated to obtain a FLER estimate. Aircraft measurements for this study were made by Scientific Aviation Incorporated, and a detailed discussion of the methods, results and uncertainty are presented in Conley et al.³⁸ It is important to note that this method differs from “mass balance methods” utilized to quantify emissions from multiple facilities over larger spatial scales.^{23,39,40}

3.3.3 ON-SITE MEASUREMENTS

On-site measurements were made by AECOM Inc., or by a combined team including AECOM and study partner personnel. On-site measurement teams moved through gathering stations from the station inlet to the station outlet and identified emission sources via OGI (FLIR[®] GF320, Opgal EyeCGas[®]). At some gathering stations, laser methane detectors were also utilized to locate emission sources (Heath Consultants RMLD-IS[®]). Detected leaks were quantified using a high-flow sampler (Bacharach Hi Flow[®]). Due to recent concerns^{41,42} about instrument accuracy and sensor transition failure, the instruments were calibrated daily according to the manufacturers’ recommendations, as specified for measuring gas with a low CH₄ concentration⁴³, even though the CH₄ fraction at gathering stations within the study area typically exceeded 0.93. Emissions observed with OGI but not measured due to personnel safety concerns or inaccessibility were documented as “observed not measured”.

Observed emissions (both measured and unmeasured) were classified by major equipment category and component type for use in study on-site estimate development. Table 3.3 provides counts of on-site measurements by equipment type, detailing valid measurements.

Table 3.3: On-site direct measurements by equipment type, detailing valid measurements, and those above and below the measurable leak rate of the high-flow sampler. Observed but not measured emission sources were categorized for simulation in the study on-site estimate (SOE).

Equipment Type	Onsite Direct Measurements				Observations	
	Valid Measurement	Above Hi-Flow Range	Below Hi-Flow Range	Total Sources Measured	Observed Not Measured	Total Sources Observed
Compressor	208	5	80	293	16	309
Dehydrator	15	-	20	35	-	35
Other	26	-	26	52	2	54
Pig Launcher/Receiver	1	-	-	1	-	1
Piping or Gas Line	25	-	15	40	-	40
Separator	25	-	27	52	-	52
Tank	9	-	2	11	8	19
Total	309	5	170	484	26	510

Dehydrator still vents were not expected to be a significant CH₄ emission source based on GRI-GLYCalc⁴⁴ simulations (an approved software program for predicting air emissions from glycol dehydrator units in 40 CFR 98.233). However, a limited number of field measurements exhibited substantially larger CH₄ emissions than predicted. Glycol dehydrators at one gathering station were equipped with passive condensers known as “BTEX Busters”, which cool the still vent exhaust stream, thereby removing entrained liquids and volatile organic compounds. The still vents on four dehydrator units were measured with the high-flow sampler at 7.6, 5.7, 5.2, and 1.2 kg/h CH₄ respectively. Four emission factors were developed by normalizing the measured emissions by the rated horsepower of operating compressor engines providing gas to each unit. Gas throughput measurements were not available for individual dehydrator units, however, throughput is directly correlated to operating horsepower as shown in Figure 3.6.

Process simulations of dehydrator still vent emissions using GRI-GLYCalc are highly sensitive to input parameters, and nullify still vent emissions if the user indicates that the simulated unit employs flash tank vapor recovery (an emission control technique). All four de-

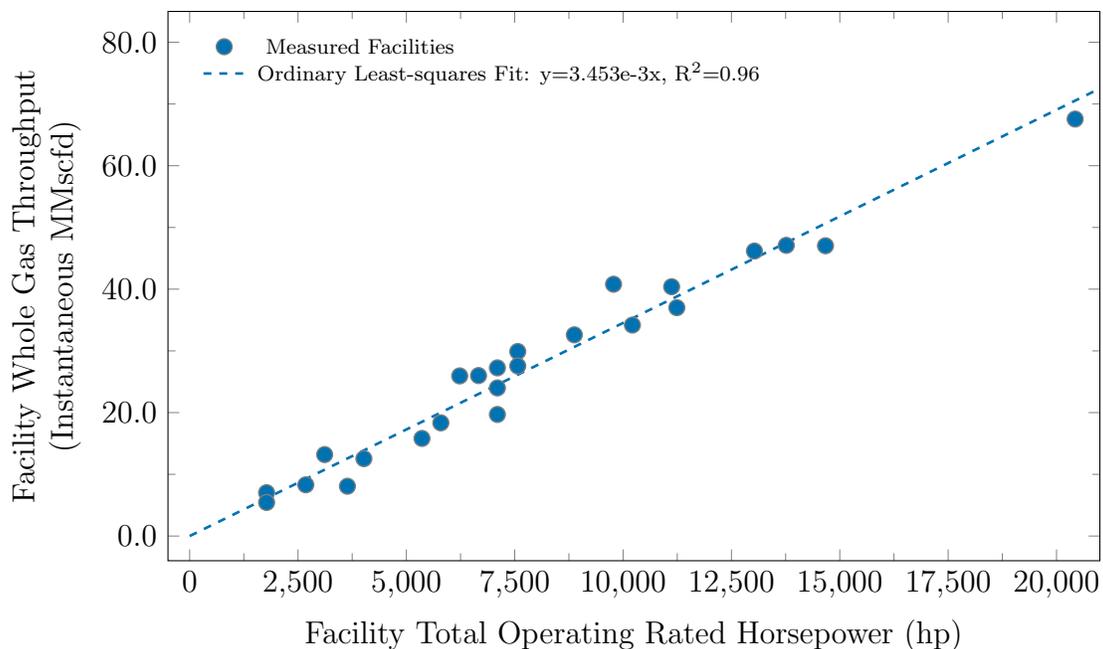


Figure 3.6: Facility total operating horsepower vs instantaneous facility total throughput, as noted by on-site observers during the field campaign.

hydrator units measured in the field campaign employed flash tank vapor recovery, but measured CH₄ emissions were larger than uncontrolled emissions predicted by GRI-GLYCalc., as shown in Table 3.4.

Table 3.4: Dehydrator still vent emission factor comparison table. Measurements of four dehydrators equipped with emission control devices exhibit significantly greater emissions than predicted by GRI-GLYCalc.

Dehydrator:	Unit 1	Unit 2	Unit 3	Unit 4	
Rated Compressor Power Input	5325	5325	5325	3115	hp
Flow from Correlation	18.4	18.4	18.4	11.2	MMscfd
Measured Regenerator Vent	7.6	5.7	5.2	1.2	kg/h
GRI-GLYCalc Controlled Regenerator Vent	0.04	0.04	0.04	0.02	kg/h
GRI-GLYCalc Uncontrolled Regenerator Vent	0.72	0.72	0.72	0.42	kg/h

3.3.4 STACK TEST MEASUREMENTS

Combustion slip (unburned fuel entrained in engine exhaust) represents a significant component of total methane emissions at gathering stations. No measurements of combustion exhaust were made during the field campaign, but study partner companies provided recent exhaust stack test data measured for state or federal regulatory compliance. These data were measured by contractors who performed stack tests in accordance with standard protocol (EPA Method 19⁴⁵, EPA Method 320⁴⁶) in the year prior to the field campaign (January to September, 2014). Stack test data were provided for 111 engines; 24 were from one engine series (Caterpillar[®] G3500, rated at ≈ 1 MW), and 87 from another (Caterpillar[®] G3600, rated at ≈ 1.3 MW). All compressor engines present at measured gathering stations belonged to one of these engine series.

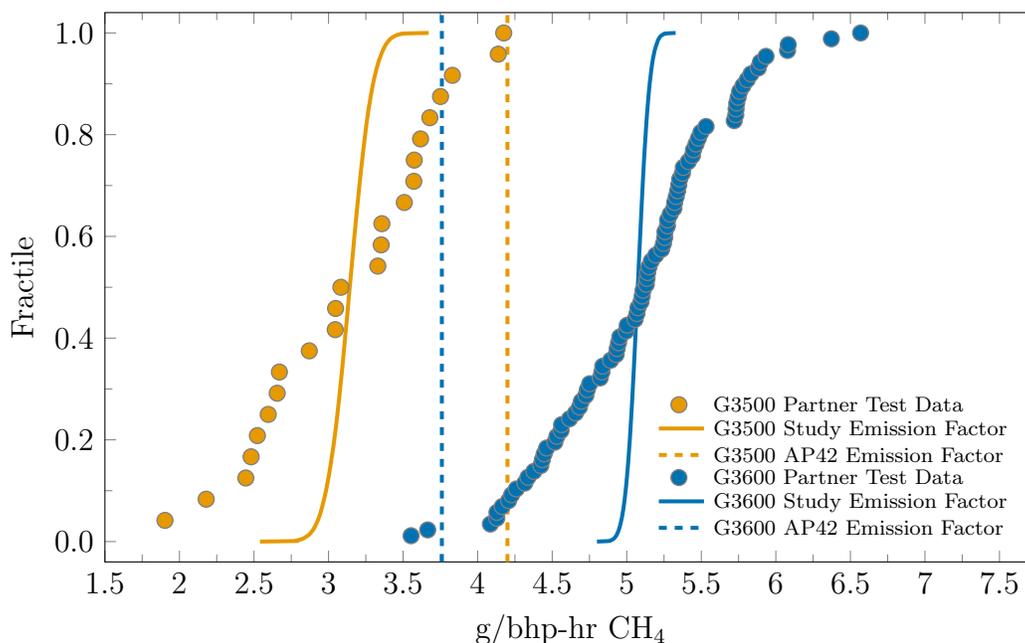


Figure 3.7: Combustion slip emission rates measured by study partners in the year prior to the field campaign (marks). Study emission factors were developed by bootstrapping measured results (solid lines) for comparison AP-42 emission factors (dashed lines).

Stack test data were normalized by the average brake horsepower of the engine during the test, as shown in Figure 3.7. Means and 95% confidence intervals developed from n-out-of-n bootstrap resampling for each engine series show statistically significant differences

in mean combustion slip (G3500: mean 3.10 g CH₄/bhp-h (\pm 0.23); G3600: mean 5.02 (\pm 0.12) g/bhp-h). This is equivalent to 4.15 (\pm 0.32) kg CH₄/h and 8.9 (\pm 0.21) kg CH₄/h respectively, for each engine series when operating at rated power. These emission rates are similar to those recently measured by Johnson et al.⁴⁷ at transmission compressor stations in the Barnett shale.

Table 3.5 compares emission factors and rates from this test data to three EPA methods: (1) greenhouse gas inventory (GHGI)⁹, (2) compilation of air pollutant emission factors (AP-42)⁴⁸, and (3) greenhouse gas reporting program (GHGRP) Subpart C (40 CFR 98.33)⁴⁹. Emission factor differences between methods (and between models of the same engine type) highlight the importance of using specific emission factors when estimating combustion slip emissions from activity data. For example, the AP-42 factor would overestimate combustion slip by 26% for measured G3500 series engines, and underestimate combustion slip from measured G3600 series engines by 34%, when assuming manufacturer rated fuel use at rated power.

Table 3.5: Combustion Slip Emission Factor Summary Table. EPA factors (GHGI, AP-42, Subpart C) assuming manufacturer rated fuel use at rated power.

	G3500		G3600	
	Factor (g/bhp-h)	Rate (kg/h)	Factor (g/bhp-h)	Rate (kg/h)
Study	3.10 (+/- 0.23)	4.15 (+/- 0.32)	5.02 (+/- 0.12)	8.9 (+/- 0.21)
GHGI	4.62	6.19	4.62	8.20
AP 42	4.2	5.63	3.76	6.67
Subpart C	7.4e-3	9.9e-3	6.6e-3	11.8e-3

3.4 MODEL DESCRIPTION

The study on-site estimate (SOE) is a comprehensive statistical estimate of CH₄ emissions from a gathering station comparable to tracer and aircraft FLER estimates. SOEs were developed from on-site direct measurements (ODMs) and engineering estimates in a Monte Carlo model. Engineering estimates were made for compressor engine crankcase vents, and glycol dehydrator (“dehydrator”) regenerator vents. Emissions from unburned methane

entrained in compressor engine exhaust (combustion slip) were estimated based on 111 recent measurements of representative engines made by measurement contractors prior to the study. The SOE in this study does not include emissions from malfunctions, maintenance, or other intended or unintended operating conditions for which on-site teams had no means to measure or accurately estimate emissions (“immeasurable sources”). Five measured facilities were excluded from method comparisons due to immeasurable sources as described in Section 3.3

To develop the SOE, direct measurements were summed for each facility and Monte Carlo methods were used to estimate emissions for known or observed sources that could not be measured. Every identified source was included, resulting in a comprehensive FLER estimate comparable to dual-tracer and aircraft measurements. On each iteration of the Monte Carlo simulation, individual SOE categories are calculated and summed as follows.

$$SOE_i = \dot{m}_{ODM,i} + \dot{m}_{SDM,i} + \dot{m}_{simcombslip,i} + \dot{m}_{simdehy,i} + \dot{m}_{simcrankcase,i} \quad (3.1)$$

On-site direct measurements (ODMs) are the sum of all measurements made by high-flow samplers at a facility during the field campaign. ODMs refer to component or device-level measurements of flanges, unions, valve stem packing, rod packing vents, connectors, pressure regulators, tank vents, open-ended lines, pneumatic devices and controllers, and other sources expected to emit within the measurable leak rate of the high-flow sampler (0.05 SCFM–8 SCFM or equivalently 0.058–9.24 kg/h)⁴³. ODMs were made with the Bacharach Hi Flow[®] sampler and individual measurement uncertainties are assumed to correspond to the instrument accuracy ($\pm 10\%$)⁴³.

For each Monte Carlo iteration, i , methane emissions from ODMs at facility j are calculated as:

$$\dot{m}_{ODM,i} = \sum_{k=1}^N f_i \cdot ODM_k \quad (3.2)$$

Where:

N is the number of on-site direct measurements made at facility j not subject to any emission rate exceptions

f_i is a factor drawn from a normal distribution to account for the high-flow sampler measurement uncertainty ($\pm 10\%$)⁴³

Simulated direct measurements (SDMs) encompass the same source categories as ODMs and provide an estimated emission rate for sources where an ODM was attempted but out of range, or would have been attempted had the source been safely accessible. SDMs do not account for immeasurable sources. Emission sources observed but not measured due to inaccessibility or personnel safety concerns were documented as “observed not measured” and were accounted for in the Monte Carlo model by re-sampling from representative ODMs made in this study. Measurements with recorded values outside the measurable leak rate range of the high-flow sampler were also accounted for in the Monte Carlo model.

For each Monte Carlo iteration, i , methane emissions from simulated direct measurements at facility j are calculated as:

$$SDM_j = \dot{m}_{obsnotmeas,i} + \dot{m}_{abovehf,i} + \dot{m}_{belowhf,i} \quad (3.3)$$

Emissions observed but not measured at facility j are sampled from the distribution of ODMs developed during this study as:

$$\dot{m}_{obsnotmeas,i} = \sum_{k=1}^N \text{draw}(\dot{m}_{ODMeqtype(k)}) \quad (3.4)$$

Where:

N is the number of observed not measured emissions sources

$\text{draw}(\dot{m}_{ODMeqtype(k)})$ indicates drawing one value at random from the distribution of measurements of equipment type k

In the event that an emission source was recorded at or above the measurable leak rate of the high-flow sampler, the measurement was removed from the ODM category, and $\dot{m}_{abovehf,i}$ was estimated by drawing a replacement emission rate from a right triangular distribution with maximum probability at the maximum measurable leak rate of the high-flow sampler (8 SCFM or 9.24 kg/h)⁴³, tapering to a minimum probability at an emission rate of 16 SCFM (18.48 kg/h), and added to the SDM category. This upper limit was chosen on the assumption that on-site measurement personnel would not attempt to measure an emission source greater than twice the measurable leak rate, and conversely that any measurement attempt would capture at least half of the emission source. OGI camera observations reinforce that this is a reasonable assumption for the instances observed during this study as shown in Figure 3.8.



Figure 3.8: Still image taken from optical gas imaging (OGI) camera footage. In this case the emission rate exceeded the measurable leak rate, and was not capturing the entire emission plume.

In the event that a measurement was observed with OGI, but registered below the measurable leak rate of the high-flow sampler (0.05 SCFM or 0.058 kg/h)⁴³, the measurement was removed from the ODM category, and $\dot{m}_{belowhf,i}$ was estimated by multiplying the measured reading by an uncertainty factor that increased from $\pm 10\%$ at the lower measurable leak rate to $\pm 100\%$ at recorded emission rate of 0 SCFM, and added to the SDM category.

Simulated Combustion Slip accounts for the CH₄ component of un-burned fuel entrained in natural gas-fired compressor engine exhaust. Study partners provided compressor engine exhaust test data from 111 engines measured in the year prior to this study; combustion slip was not measured in this study. This sample represents an estimated one fourth of all

gathering compressor engines within the study area. Tests were performed on engines located within the study area by measurement contractors using standard protocol (EPA Method 19⁴⁵, EPA Method 320⁴⁶). Test data were provided for 24 Caterpillar[®] G3500 series engines and 87 Caterpillar[®] G3600 series engines. These engine series represent approximately 93% of all gathering compressor engines within the study area; all compressor engines at measured gathering stations belonged to one of these engine series. This ensured the applicability of exhaust test data and resulted in improved combustion slip estimates relative to compiled emission factors such as EPA AP-42⁴⁸. No uncertainty is provided for individual engine measurements; uncertainty is developed in the Monte Carlo model from the variation in measured combustion slip within an engine series.

For each Monte Carlo iteration, i , combustion slip methane emissions for facility j are calculated as:

$$\dot{m}_{simcombslip,i} = \sum_{k=1}^{N_{op}} \text{draw}(EF_{series(k)}) \cdot \text{draw}(Load_k) \cdot RatedHP_k \quad (3.5)$$

Where:

N_{op} represents the count of compressor engines operating on-site during the measurement

$\text{draw}(EF_{series(k)})$ indicates drawing one emission factor value at random from the distribution of emission factors for the same engine series as engine k

$\text{draw}(Load_k)$ indicates drawing a fractional load at random from the distribution of operating loads observed during the filed campaign, and applying it to engine k

$RatedHP_k$ is the rated power output of engine k

Simulated Dehydrator Regenerator Vents account for CH₄ emissions from dehydrator regenerator vents. All dehydrators at measured gathering stations employed flash tank vapory recovery systems, an emission control technique. Methane emissions from dehydrator

regenerator vents were calculated in the Monte Carlo model using the emission factor for dehydrators with flash tank vapor recovery from a 1996 GRI study⁵⁰ (0.003 (-52%/+102%) kg/h CH₄ per MMscf per day of gas processed).

Simulated Crankcase Vents account for CH₄ vented from compressor engine crankcase vents because of imperfect piston ring sealing. Crankcase vents on compressor engines were not measured in this study, but were simulated based on a Caterpillar[®] crankcase ventilation application guide⁵¹, which states that crankcase hydrocarbon emissions are normally 3% of total hydrocarbon exhaust emissions at engine mid-life, but could be as high as 20% due to engine wear. Crankcase emissions were calculated in the Monte Carlo model by multiplying combustion slip emissions by a factor drawn at random from a normal distribution (mean 3%, assumed standard deviation 2%).

Recent measurements of dehydrator regenerator (this study) and crankcase vents⁴⁷ may indicate that these categories are conservatively estimated in SOEs. Direct measurements of regenerator vents on four dehydrators equipped with flash tanks and regenerator control devices were made in this study. Measurements were normalized by the gas throughput of each dehydrator, resulting in emission factors of 0.11, 0.28, 0.31, and 0.41 kg/h CH₄ per MMscf per day of gas processed (see Table 3.4). Emission factors based on these four measured units are one to two orders of magnitude greater than the Gas Research Institute (GRI) study emission factor for dehydrators with flash tanks and overlap or exceed those provided in the GRI study for dehydrators without flash tanks (0.14 (-50%/+101%) kg/h CH₄ per MMscf per day of gas processed). Flash tanks are an emission control technique that can reduce methane emissions by 90%⁵²; the regenerator vent control devices on measured units do not affect methane emissions. GRI-GLYCalc nullifies flash tank emissions when the user indicates the presence of flash tanks on a simulated unit. Johnson et al.⁴⁷ measured crankcase vent methane emissions and combustion slip on Caterpillar[®] 3500 and 3600 series compressor engines and found crankcase vent emissions were 14.4% of combustion slip on average (range 7%–22%). Engine wear, which is a primary cause of increased crankcase

vent emissions, cannot be readily deduced for engines at gathering stations in this study. Alternate method comparisons of tracer and aircraft to SOEs developed using these recent measurements of crankcase vents and dehydrator regenerator vents are provided in Appendix B.1.

Methane emissions from acid gas removal (AGR) units are not included in SOEs. Two gathering stations with AGR units were measured in the study but were excluded from method comparisons because of incomplete measurement and immeasurable sources, respectively, as described in Section 3.3. No measurements of AGR unit reboiler vents were made in this study.

3.5 RESULTS

Comparisons are made between tracer and SOE, or aircraft and SOE. Tracer and SOE are compared at facilities measured concurrently in the absence of immeasurable sources. Aircraft and SOE are compared at facilities measured contemporaneously assuming the absence of immeasurable sources. Method comparisons were performed using the approaches of Bland and Altman⁵³, and Neri et al.⁵⁴ The approach of Bland and Altman (“difference plot”) is generally accepted as the appropriate technique for analyzing method comparison studies⁵⁵. It indicates the presence or absence of bias between methods, and provides an estimate of expected agreement between methods based on the sample population. The approach of Neri et al. is a variance-weighted least-squares (VWLS) regression that minimizes the orthogonal distance between measurement data points and the line of best-fit, considering the error in both x and y data (see Appendix B.2). Additionally, a bootstrap⁵⁶ was performed to estimate a 95% confidence interval on the VWLS regression slope. Ten-thousand new input datasets were constructed from SOE distributions output by the Monte Carlo model and normal distributions created from tracer or aircraft measurements and associated uncertainty. VWLS fits were performed on each re-sampled dataset; the 2.5th and

97.5th percentile of re-sampled VWLS regression slopes provide a 95% confidence interval on the original regression slope.

Ordinary least-squares regression is inappropriate because the regression slope depends on the choice of independent variable and both measurements are error affected. The slopes of ordinary least-squares regressions from either choice of independent variable may not bound the slope of the orthogonal regression.⁵⁷ Orthogonal regression is required to predict results from one measurement method when measurements were made with another measurement method and both measurement methods are error affected.⁵⁸

Simulated Combustion Slip was the largest source category and contributed 78% to the cumulative SOE for the 17 facilities included in method comparisons shown in Figure 3.9. *ODMs* contributed 15%, *SDMs* contributed 5%, *Simulated Crankcase Vents* contributed 2%, and *Simulated Dehydrator Regenerator Vents* contributed less than 1% to the cumulative SOE. For each measurement method, 95% confidence intervals indicate that the method would produce a FLER within the interval 95% of the time. We consider methods with overlapping confidence intervals to agree. Tracer and SOE 95% confidence intervals overlap at 11 out of 14 facilities, while aircraft and SOE confidence intervals overlap at three out of six facilities. However, if each method provides an approximate quantification of the true FLER, measurements with greater uncertainty are more likely to agree.

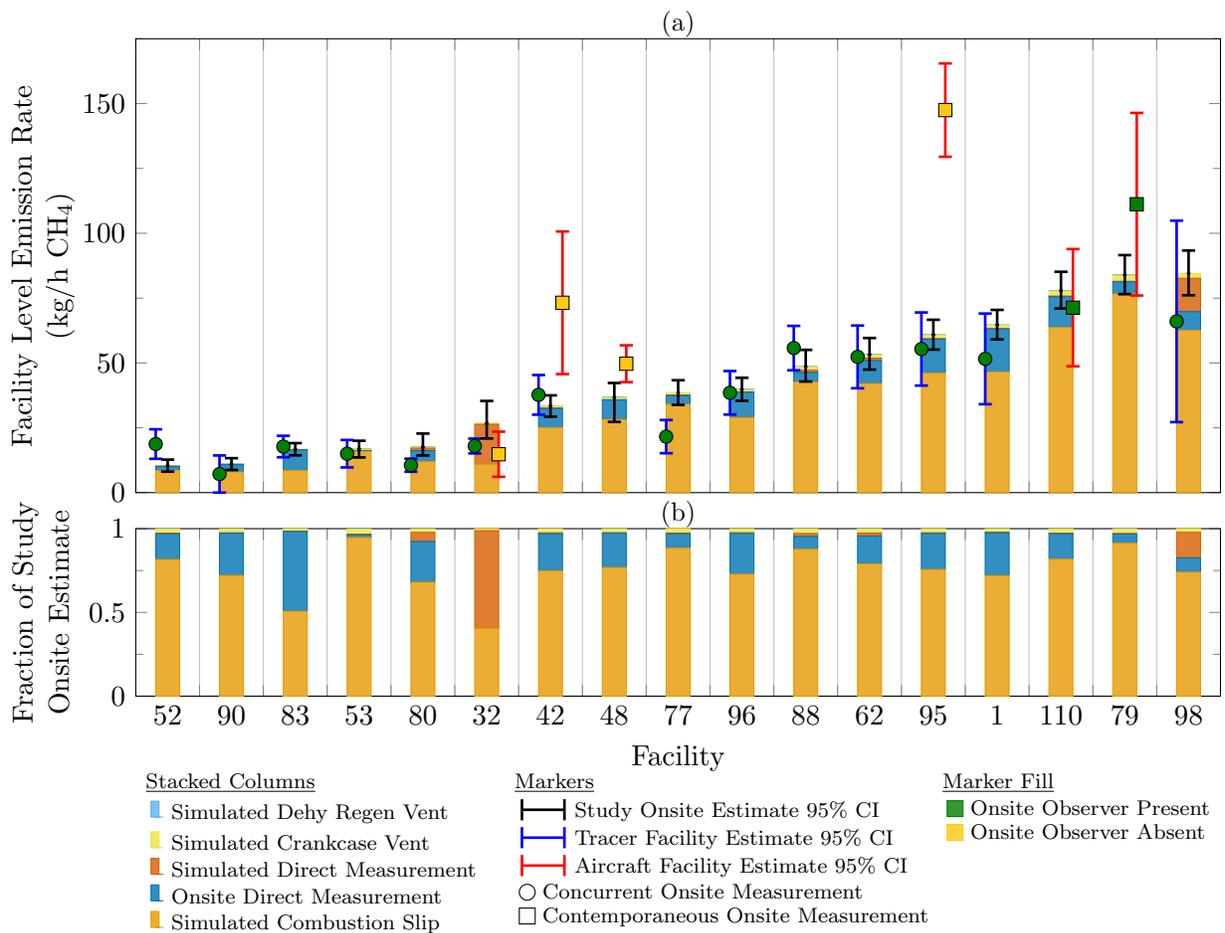


Figure 3.9: Facility-level CH₄ emission rate summary at all facilities included in method comparisons. Study on-site estimates (SOE) are the sum of on-site direct measurements plus engineering estimates for unmeasured sources (stacked columns, black error bars). Tracer (left mark, blue error bars) and aircraft (right mark, red error bars) are overlaid at facilities where these measurements were compared to SOEs. Marker shape and fill indicate same/different day and the presence/absence of on-site observers, which influence the comparability of measurements. Bottom panel illustrates the fraction of the SOE contributed by each component; combustion slip contributes more than half of emissions at 16 of 17 facilities and accounts for three quarters of cumulative SOE emissions for these 17 facilities.

3.5.1 TRACER FACILITY ESTIMATE AND STUDY ON-SITE ESTIMATE COMPARISON

When compared in aggregate by difference plot and variance-weighted least-squares regressions, tracer predicts lower FLER than SOE for 14 concurrently-measured gathering stations at the 95% confidence level. In Figure 3.10 the difference of tracer and SOE is plotted against the uncertainty weighted mean of tracer and SOE. The mean of differences (termed “bias”) is -4.9 kg/h, indicating that tracer predicts lower FLER than SOE. A paired t-test is used to determine if the bias is significant. The shaded area in Figure 3.10 highlights the 95% confidence interval on bias. The confidence interval does not include $x = 0$, which indicates that the bias is statically significant at the 95% confidence level.

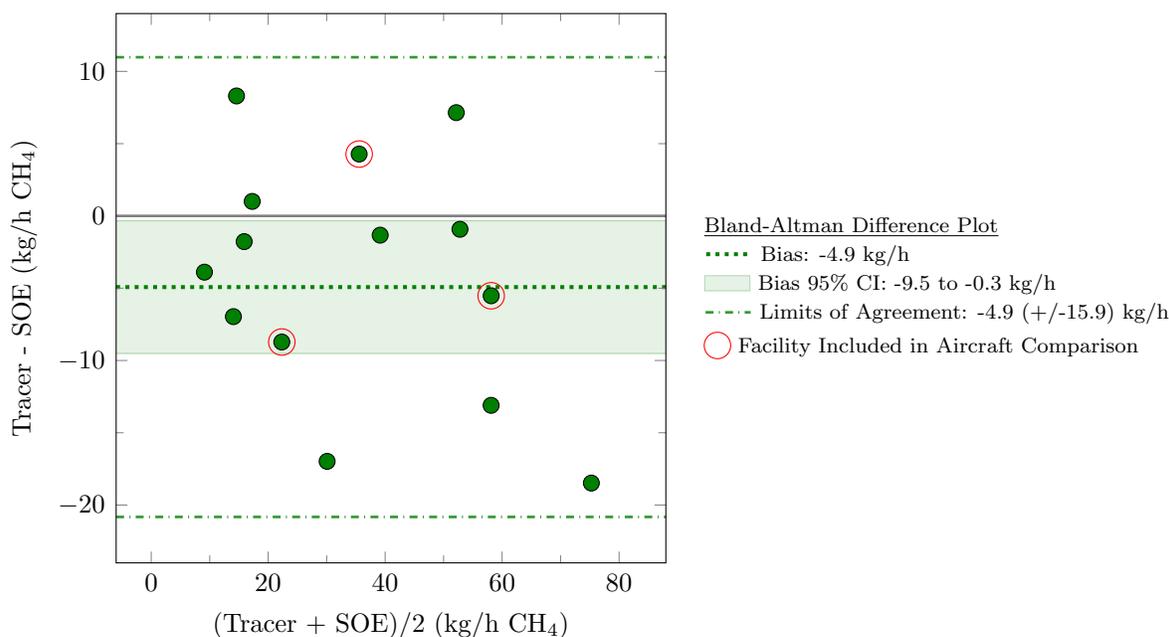


Figure 3.10: Tracer predicts lower facility-level CH₄ emission rates than study on-site estimates at the 95% confidence level using difference plot.

In Figure 3.11 a VWLS regression (dashed line) is performed on tracer and SOE. The slope of the regression (tracer = 0.91·SOE, $R^2 = 0.89$) is less than unity, indicating that tracer predicts lower FLER than SOE. The 95% confidence interval (shaded region) on the regression slope (tracer = 0.83·SOE to tracer = 0.99·SOE) does not include the line of equality ($y = x$), indicating that tracer predicts lower FLER than SOE at the 95% confidence

level. A fundamental assumption of the tracer method is that tracer gases released at a facility undergo the same dispersion as the target analyte emitted from the facility (CH₄ in this case). Buoyant combustion plumes may violate this co-dispersion assumption and result in a low-biased FLER. Roscioli et al.¹⁶ estimated worst-case recovery of combustion slip for several scenarios using Gaussian dispersion modeling with Briggs plume rise equations. They found that tracer may not recover up to 50% of the combustion slip plume when downwind measurements are made at distances of less than 1000 m. Recovery improves with increasing downwind distance. Yacovitch et al.³⁷ found no evidence of plume rise at gathering stations measured in this study by comparing plume emissions to downwind measurement distance. However, this finding is not absolutely conclusive because downwind measurement distance varied little since it was dictated by the presence of roads (see Figure S14 in Yacovitch et al.³⁷). Assessing plume recovery by releasing tracer gases directly into compressor engine exhaust stacks in conjunction with tracers placed on the ground in typical locations, as done in Lamb²⁰, may be a worthwhile addition to future tracer measurements at facilities with CH₄ emissions entrained in elevated, buoyant plumes.

Appendix B.1 discusses an alternate method comparison where SOEs were developed based on recent direct measurements of crankcase vents⁴⁷ and dehydrator regenerator vents (this study). The results of this comparison also indicate that tracer predicts lower FLER than SOE (regression of tracer to SOE = 0.76 (95% CI = 0.69 to 0.83), $R^2 = 0.92$).

Table 3.6 details the individual source contributions to the cumulative SOE for the 14 gathering stations included in TFE/SOE method comparisons. Simulated combustion slip accounts for 391.5 kg/h, or 75% of the cumulative SOE. On-site direct measurements account for 86.9 kg/h or 17% of the cumulative SOE, while simulated direct measurements account for 32.1 kg/h or 6% of the cumulative SOE. Compressor engine crankcase vents contribute 11.7 kg/h or 2% of the cumulative SOE, while simulated dehydrator regenerator vents contribute 0.8 kg/h or less than 1% to the cumulative SOE. Rod packing vents and pressure relief valves were the only ODM categories where measurements exceeded the

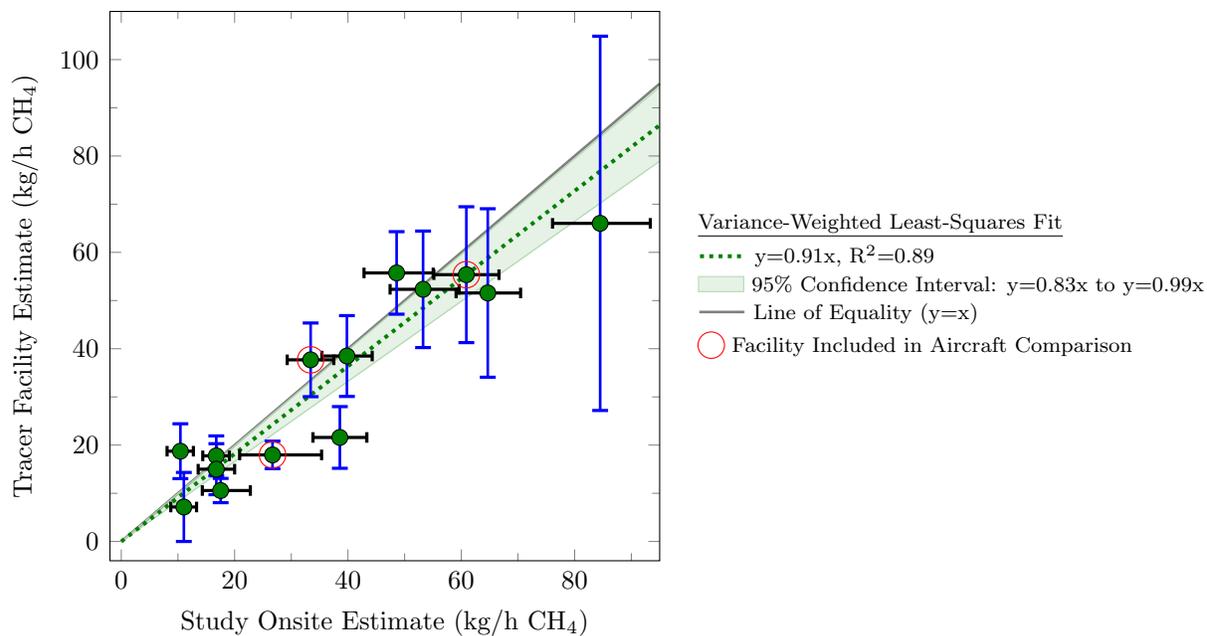


Figure 3.11: Tracer predicts lower facility-level CH_4 emission rates than study on-site estimates at the 95% confidence level using variance-weighted least-squares regressions.

range of the high-flow sampler. Observed but not measured sources appear most often on compressors, where high temperatures, complex piping, and rotating equipment make direct measurements challenging.

Table 3.6: Comparison of cumulative CH₄ emission rates for 14 gatherings stations included in the TFE/SOE comparison, showing all categories contributing to the SOE. Combustion slip is the largest contributor to the SOE.

CH ₄ Emission Source	Onsite Direct Measurements	Simulated Direct Measurements			Simulated
		Above Hi-Flow Range	Below Hi-Flow Range	Observed Not Measured	
Compressor Units	42.6	-	0.5	5.6	-
Pressure Relief Valves	3.4	12.3	-	-	-
Rod Packing Vents	27.3	12.3	0.2	-	-
Dehydrator Units	2.2	-	0.1	-	-
Other	4.5	-	0.3	0.5	-
Pig Launchers/Receivers	0.1	-	-	-	-
Piping or Gas Lines	0.4	-	0.2	-	-
Separators	1.7	-	-	-	-
Tanks	4.8	-	0.0	-	-
Combustion Slip	-	-	-	-	391.5
Crankcase Vents	-	-	-	-	11.7
Dehydrator Regenerator Vents	-	-	-	-	0.8
Cumulative Study Onsite Estimate	86.9	24.7	1.3	6.1	522.9 kg/h
Cumulative Tracer Facility Estimate					466.0 kg/h

3.5.2 AIRCRAFT FACILITY ESTIMATE AND STUDY ON-SITE ESTIMATE COMPARISON

Aircraft and SOE are compared at 6 facilities measured contemporaneously by difference plot and VWLS regression. Confidence in this comparison is reduced relative to the tracer and SOE comparison, because measurements were made between 1 and 22 days apart. Additionally, the absence of immeasurable sources could not be confirmed when on-site observers were absent during measurements. Two facilities were measured with on-site observers present during both measurements, and four facilities were measured without observers present during aircraft measurements.

Aircraft predicts higher FLER than SOE when compared by difference plot and VWLS regression. When compared by difference plot, aircraft is biased high relative to SOE (32.4 kg/h), as shown in Figure 3.12 . However, the bias is not statistically significant because the 95% confidence interval includes $x = 0$; however, the bias is significant at the 90% confidence level.

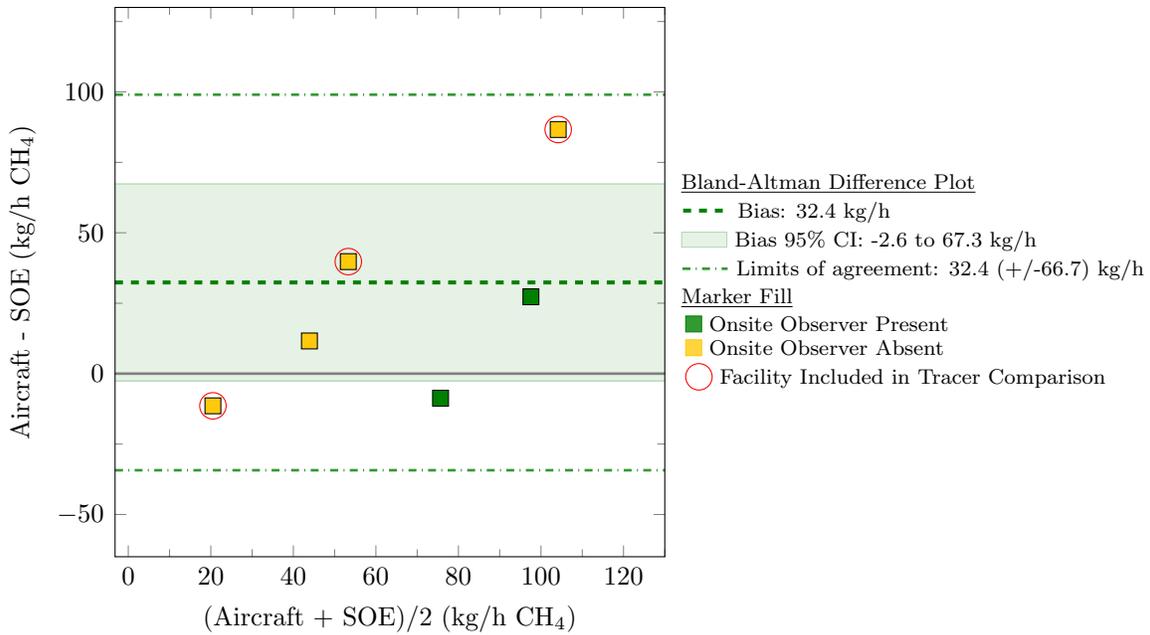


Figure 3.12: Aircraft Facility Estimate vs Study On-site Estimate Dehy In Difference Plot

In Figure 3.13, a VWLS regression (dashed line) is performed on aircraft and SOE. The slope of the regression (aircraft = 1.49·SOE, $R^2 = 0.53$) is greater than unity, indicating that aircraft predicts higher FLER than SOE. The 95% confidence interval (shaded region) on the regression slope (aircraft=1.32·SOE to aircraft=1.67·SOE) does not include the line of equality ($y = x$), indicating that aircraft predicts higher FLER than SOE at the 95% confidence level.

The observed bias in the aircraft and SOE method comparisons may be partly explained by the inability of aircraft to partition emissions from facilities within (or very near) the flight boundary. The loops flown by the aircraft covered significant area beyond the target gathering station and often included other facilities due to the geographic density of facilities in an active gas field. Post-campaign analysis of activity data indicated that emissions from nearby wells were included in aircraft measurements at least twice during the field campaign (see Section 3.3). These field observations suggest that aircraft facility measurements should be utilized with caution when emissions from nearby facilities may confound results. Aircraft facility measurements may not suffer this limitation when measuring emissions from facilities

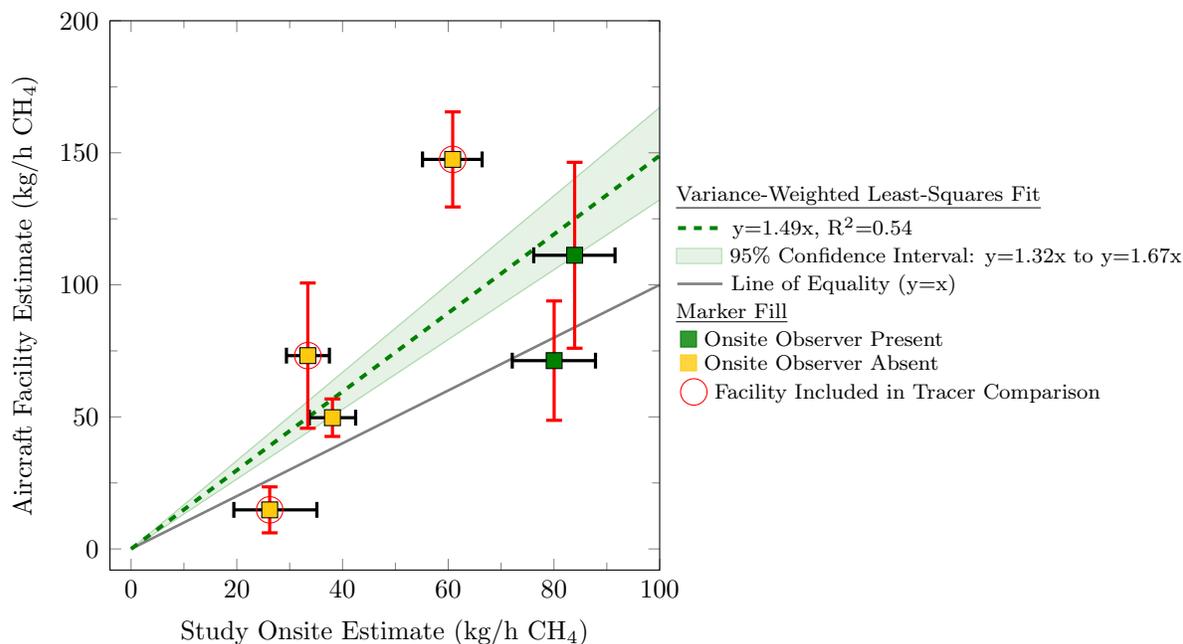


Figure 3.13: Aircraft Facility Estimate vs Study On-site Estimate Variance-Weighted Least-Squares Plot

without interfering sources, for example, well-isolated transmission and storage facilities, power plants, and landfills.

Additionally, the lowest flight altitude is limited by safety and regulations. Calculation of emission fluxes therefore must extrapolate emission rates from the lowest loop to the ground level. This is generally the largest source of method uncertainty.³⁸ Equipping the aircraft with tracer gas measurement capabilities and using tracer release gases could help to isolate emissions originating from the target facility from other nearby facilities, and may help quantify the effects of extrapolating calculated fluxes to ground level by evaluating tracer recovery rates under real field conditions.

Appendix B.1 discusses an alternate method comparison where SOEs were developed based on recent field campaigns where crankcase vents⁴⁷ and dehydrator regenerator vents (this study) were measured directly. The results of this comparison also indicate that aircraft predicts higher FLER than SOE (regression of aircraft to SOE = 1.22 (95% CI = 1.08 to 1.38), $R^2 = 0.54$).

Table 3.7 details the individual source contributions to the cumulative SOE for the 6 gathering stations included in AFE/SOE method comparisons. Simulated combustion slip accounts for 253.3 kg/h, or 79% of the cumulative SOE. On-site direct measurements account for 44.5 kg/h or 14% of the cumulative SOE, while simulated direct measurements account for 16.4 kg/h or 5% of the cumulative SOE. Compressor engine crankcase vents contribute 7.6 kg/h or 2% of the cumulative SOE, while simulated dehydrator regenerator vents contribute 0.7 kg/h or less than 1% to the cumulative SOE.

Table 3.7: Emissions by category in the SOE for the 6 gathering station included in the AFE SOE method comparison.

CH ₄ Emission Source	Onsite Direct Measurements	Simulated Direct Measurements			Simulated
		Above Hi-Flow Range	Below Hi-Flow Range	Observed Not Measured	
Compressor Units	11.9	-	0.3	3.2	-
Pressure Relief Valves	0.6	-	-	-	-
Rod Packing Vents	17.8	12.3	0.1	-	-
Dehydrator Units	2.2	-	0.2	-	-
Other	0.6	-	0.0	-	-
Pig Launchers/Receivers	-	-	-	-	-
Piping or Gas Lines	4.6	-	0.2	-	-
Separators	2.9	-	0.2	-	-
Tanks	3.9	-	0.0	-	-
Combustion Slip	-	-	-	-	253.3
Crankcase Vents	-	-	-	-	7.6
Dehydrator Regenerator Vents	-	-	-	-	0.7
Cumulative Study Onsite Estimate	44.5	12.3	0.9	3.2	322.6 kg/h
Cumulative Aircraft Facility Estimate					467.7 kg/h

Recall that facilities with immeasurable sources were excluded from method comparisons. Therefore, cumulative SOE contributions are representative of “normally operating” gathering stations within the study area, and are not representative of all gathering stations within, or outside of, the study area. The relative contribution of source categories would change if immeasurable sources were included. For example, emission rates from tanks at two gathering stations were much greater than the measurement capability of high-flow samplers used by on-site teams. These facilities are not included in method comparisons; each occurrence

is described in Section 3.3. An SOE was calculated for all sources except tanks at these facilities, and this result was compared to tracer and aircraft measurements to estimate the magnitude of tank emissions.

At gathering station 33, the tracer FLER (182 kg/h) was four times greater than the SOE (42 kg/h). Tank venting emissions were estimated by subtracting the SOE from the tracer measurement, because the SOE captured all emissions except those emanating from the tank, leading to estimated tank venting emissions of 140 kg/h. Gathering station 121 was measured by aircraft three times (October 2, 3, and 14, 2015 resulting in aircraft provided FLER estimates of 276 (± 99 kg/h), 676 (± 119 kg/h), and 739 (± 107 kg/h). Tracer and on-site teams visited station 121 and tracer estimated 606 (± 278 kg/h) venting from a produced water tank (see Section 3.3). Tank venting emissions of a similar magnitude were also observed in Mitchell et al.¹⁴

Dual-tracer release measurements of gathering stations were made previously in a national study by Mitchell et al.¹⁴; however, no on-site or aircraft measurements were made. These measurements were used by Marchese¹⁵ et al. to develop a national estimate of CH₄ emissions from gathering stations, which indicates an average emission rate of 53,066 scf/day (43 kg/h) per station. This result was used in the 2016 EPA GHGI⁹ as the per-facility emission rate for gathering stations.

The influence of tank venting from these two stations on the average FLER for 31 gathering stations measured by tracer in this study is shown in Figure 3.14. The data series ‘Tracer Excluding Significant Tank Venting’ uses the SOE at these two facilities to account for emissions from those facilities other than tank venting. The data series ‘Tracer’ uses the tracer measurement from station 33, and adds the tank venting emission estimate made by tracer at station 121 to the SOE for station 121 to estimate a complete FLER. The average FLER for gathering stations measured by tracer in this study, excluding emissions from significant tank venting, is 50.4 kg/h, a 17% increase over the GHGI per-facility estimate. The average FLER for gathering stations measured by tracer in this study, including emis-

sions from significant tank venting, is 74.5 kg/h, a 73% increase over the GHGI per-facility estimate.

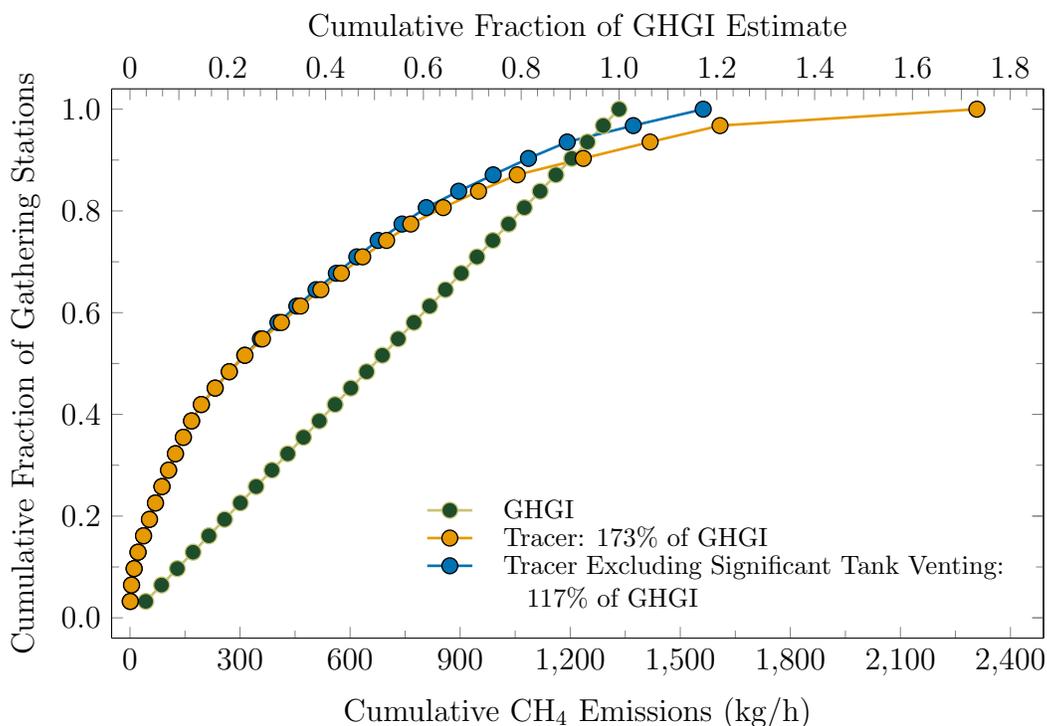


Figure 3.14: Cumulative fraction of tracer measurements compared to GHGI, both including and excluding tank venting emissions observed at two gathering stations.

At another gathering station, a compressor engine was shut down and several attempts were required to restart it. The tracer team saw increased and highly variable emissions from the facility during the restart; on-site teams had no means quantify these emissions.

3.6 CONCLUSIONS

This study provides the first contemporaneous, and in many cases, concurrent comparisons of facility-level CH₄ emissions from gathering stations utilizing both direct and atmospheric (downwind) measurement methods. SOEs were developed in a Monte Carlo model from on-site direct measurements and engineering estimates. Combustion slip contributed 78% to the cumulative SOE and was modeled using exhaust test data from 111 engines measured in the study area in the year prior to this study (January to September, 2014).

Two engine series were tested, and all engines at measured gathering stations belonged to one of them. The quality and specificity of this test data improved the accuracy of study on-site estimates by providing more accurate quantification of combustion slip than would have been possible using compiled emission factors (see Section 3.3). The clarity of method comparisons was improved by on-site observers who documented maintenance, episodic, and malfunction events during measurements. This information was used along with study partner operational data to identify facilities for exclusion from method comparisons where measurement methods were affected unequally and bias would result.

This unique combination of circumstances greatly reduced the uncertainty for pair-wise comparison of the two primary methods for estimating CH₄ emissions from larger natural gas facilities – bottom-up estimates based on detailed, device-level, on-site measurements and top-down estimates based on downwind measurements of tracer and target-gas concentrations. The reduced uncertainties, in turn, provided a high confidence indication that while tracer and SOE show strong correlation ($r = 0.91$), tracer methods may under-estimate emissions from this type of facility. As suggested by previous studies, one hypothesis for FLER underestimation by tracer relative to SOE is that CH₄ entrained in buoyant compressor engine exhaust plumes released from stacks above building height may not be fully recovered by tracer when releasing tracer gases and sampling at ground level only. Future tracer studies at gathering stations should test for co-dispersion by releasing tracer gases from exhaust stacks.

Additional advancements are likely needed for aircraft-based measurements of individual facilities to be successful in areas where it is difficult to isolate the target facility from nearby sources. Advancements could include tracer gases released at the target facility to distinguish target facility emissions from nearby sources, and improved methods for estimating mass flux below the lowest level flown by the aircraft. Study methods and results also highlighted additional areas of interest to future field campaigns. At two gathering stations, produced water tanks were observed emitting at rates several times that of all other sources at the

facility combined (see Section 3.3). Knowledge of the frequency and duration of these types of emission sources would provide a better understanding of their contribution to overall methane emissions from gathering stations. Direct measurements of regenerator vents on four dehydrators equipped with flash tanks and regenerator control devices were made in this study. All four showed emission rates greater than predicted by modeling software for dehydrators both with and without flash tanks, indicating the need for further empirical characterization of this source and validation of software used to predict methane emissions.

CHAPTER 4

RECONCILING TOP-DOWN AND BOTTOM-UP METHANE EMISSION ESTIMATES FROM NATURAL GAS OPERATIONS IN THE FAYETTEVILLE SHALE³

4.1 INTRODUCTION

Recent studies^{35,59–61} have quantified methane emissions at various scales using both direct component-level measurements and indirect measurements of atmospheric mixing ratio enhancements. Estimates of regional methane fluxes from measurements made during “top-down” aircraft mass balance flights^{39,62–64} are typically larger than estimates developed from “bottom-up” inventories^{39,59,62–65}, but have been found to be similar^{39,66}. Here, we consider aircraft mass balance flights that estimate methane emissions from a region using a “box model”. Briefly, transects are flown up-wind and down-wind of the study region (see Figure 4.1), and a flux is calculated for both the upwind and downwind sides of the box defined by the transect flight path, the ground, and top of the planetary boundary layer. Methane emissions originating from within the box are given by the difference between downwind and upwind fluxes. For a detailed description of the top-down mass balance flight methods and results used in this study, see Schwietzke et al.²³ Bottom-up inventories are developed by multiplying emission factors by activity factors. Knowledge of activity factors (i.e. counts of emitters) and emission factors for each type of emitter present within the box allows the creation of a bottom-up inventory comparable to the top-down aircraft mass balance flight.

In practice, it is challenging to fairly compare top-down and bottom-up estimates. Top-down measurements made in aircraft mass balance flights capture emission snapshots in real-time and may not fully characterize temporal variability in emissions.^{59,62} Top-down flights may also have difficulty with source attribution⁶³, particularly in basins where the methane fraction in whole gas varies substantially. Bottom-up estimates may be based on

³This chapter forms the basis of a manuscript in preparation: Vaughn, T. L., et al. Reconciling Top-Down and Bottom-Up Methane Emission Estimates from Natural Gas Infrastructure in the Fayetteville Shale. **2017**.

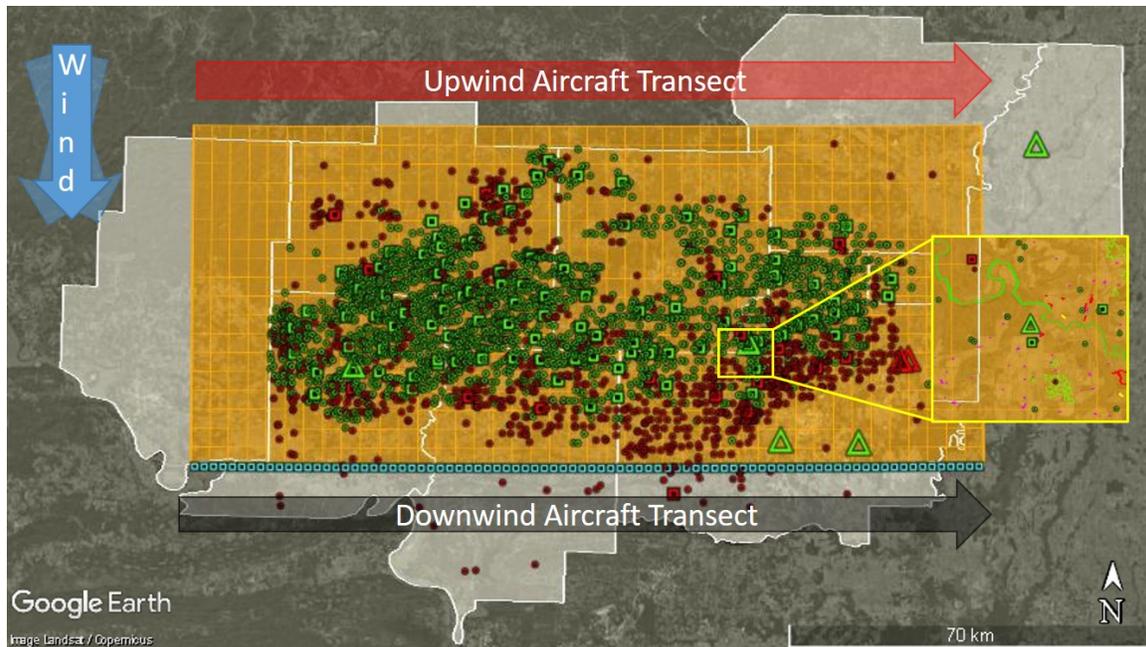


Figure 4.1: Study area overview showing the spatial extent and grid cells used in the bottom-up ground level area estimate (GLAE) and aircraft mass balance flight transects. The southern edge of the study area grid shows the receptor points use for simulated downwind transects. Inset showing spatially explicit production, gathering, and transmission facilities, distribution service regions, and wetland areas. Production facilities (circles), gathering stations (squares) and transmission stations (triangles) operated by study partners are shown in green. Non-study partner facilities are shown in red.

disaggregated national, or other, inventories, which require assumptions for apportionment and make inter-comparisons of different studies difficult without further analysis⁶⁷. Inventories have been shown to have unrepresentative emission factors¹⁷, outdated activity data¹⁸, and have in the past omitted certain source categories¹⁵. All recent studies have observed skewed distributions of both site based and component level inventories³⁴, where a small population of components or facilities contribute a disproportionate fraction of the overall emissions. The contribution of these less-common “super-emitters” may not be accurately captured in emission factors used in inventory development⁶⁸. Some sources may not even be considered in bottom-up inventories⁶⁹. For these reasons, many have suggested that top-down and bottom-up estimates be made concurrently^{28,59,63,70}.

A critical assumption employed in prior top-down to bottom-up comparisons is that activity data derived from annual averages is representative of activity in the study area

during the time of the mass balance flight. For certain source categories, this is a reasonable assumption. Emissions from malfunctions and devices actuated automatically by process parameters can happen at any time, without human intervention. For other categories, emissions are driven by human activity and occur more frequently during working hours. For example, manual liquid unloadings (MLUs) and blow-downs related to maintenance activities may produce high emission rates over short durations and are triggered by human operators during their daily work. If there are 260 working days in a year and working days are 8 hours long, daily peak emissions from these source categories could be shifted by a factor of four from an annual average. Therefore, measurement results from a mid-day mass balance flight may accurately quantify methane emissions from a study region but substantially overestimate annualized emissions from some sources.

Here, we examine the results from a controlled top-down and bottom-up study using detailed activity data and contemporaneous measurements of emissions from all segments of the natural gas supply chain present in the study area. Methane emissions from non-study partner facilities and non-oil and gas sources are also considered, and are modeled at the finest spatial granularity enabled by source data (see Section 4.2). Spatially and temporally resolved activity data was used with methane emission measurements from this and prior studies in a Monte Carlo model to develop bottom-up ground-level area estimates (GLAEs). Daily average wind speed and direction were utilized to model transport delay from emission sources to aircraft down-wind transect locations.

Herein, we first present hourly GLAEs for both days with corresponding aircraft mass balance flights. Then, we present GLAEs for the mid-day period aligned with the down-wind transect times of aircraft mass balance flights, which enables direct comparison of top-down and bottom-up results. Finally, we compare longitudinal methane emission rate profiles to those presented in Schwietzke et al.²³ by propagating GLAEs results downwind using a Gaussian dispersion model.

4.2 METHODS: MODEL DESCRIPTION

Top-down and bottom-up reconciliation was attempted by constructing a comprehensive spatio-temporal Monte Carlo model of CH_4 emission sources. The spatial boundary of the model is defined by the “study area”, which was dictated by the flight boundaries during the aircraft mass balance flights, shown in green and black in Figure 4.2; the temporal period spans the two days corresponding to aircraft mass balance flights (October 1st and 2nd, 2015), and simulates CH_4 emissions from a variety of sources on each hour of those days (“study period”). The model has two spatial elements: The GLAE, and the downwind transect estimate. The GLAE is compared to totals estimated by the aircraft mass balance, and the downwind transect estimate is compared to longitudinal emission rate profiles predicted by the aircraft. Grid cells for the GLAE are based on 0.04° longitude increments, with two downwind transect grid points per north/south column of the GLAE grid, which serve as receptors for the Gaussian dispersion model.

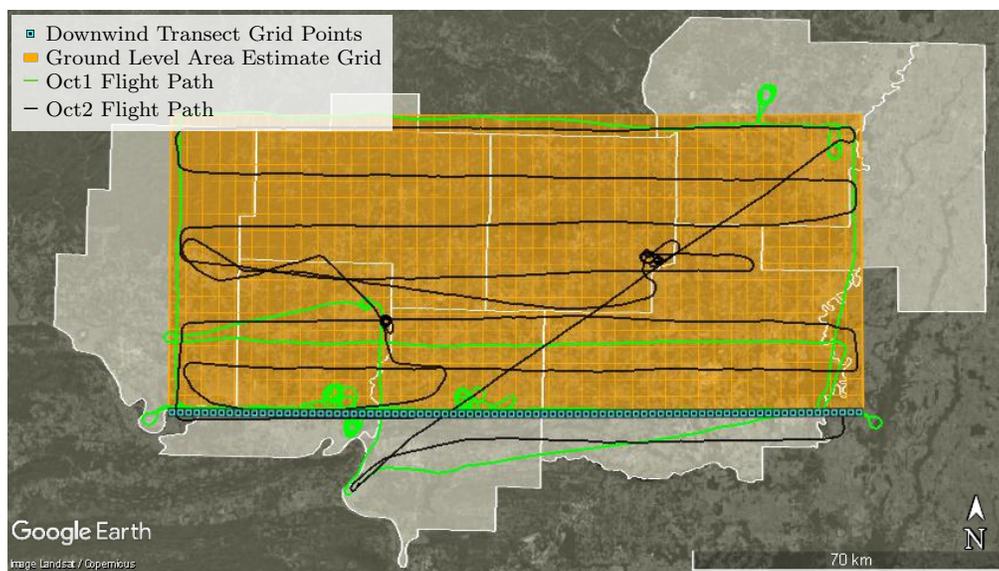


Figure 4.2: The study area is defined by the aircraft mass balance flight boundaries (orange shading), and is divided into grid cells for the spatially and temporally resolved ground level area estimate. Downwind transect grid points act as receptor points for a Gaussian dispersion model which consider the centroid of each grid cell a point source for all CH_4 emissions within the grid cell. These results can be compared directly to emission rate estimates vs longitude estimated by the aircraft.

On each iteration of the Monte Carlo model, calculated CH₄ emissions are assigned to grid cells, and are propagated downwind according to Gaussian dispersion theory based on the prevailing wind speed, wind direction, and atmospheric stability class during the aircraft mass balance flights. Emissions are calculated for contributing source categories as described in the following sections.

Table 4.1: Source categories that contribute to modeled CH₄ emission rates predicted by the ground level area estimate (GLAE).

Model Categories	
Oil and Gas	Non-Oil and Gas
Production	Livestock
Gathering	Geologic Seeps
Transmission	Wetlands
Distribution	GHGRP Facilities
	Landfills
	Rice Cultivation
	Wastewater Treatment

4.3 OIL AND GAS RELATED METHANE SOURCES

4.3.1 PRODUCTION

Emissions from the production sector are modeled based on the SOE of Bell et al.⁷¹, a comprehensive FLER estimate. Modeled emissions are categorized as shown in Table 4.2. The GLAE model modifies the calculation of manual liquid unloadings described in Bell et al. to account for transport delay from the location of the unloading to the aircraft downwind transect location.

Manual Unloadings includes emissions from vented MLUs initiated by workers as a part of normal operations within the study area. MLUs are modeled based on study partner provided activity data and emission rates from a study of liquid unloadings at U.S. natural gas production facilities by Allen et al.⁷² Study partners provided spatially and temporally explicit activity data for unloadings at individual wells, including the start times and durations of unloading events. The GLAE utilizes emission rates for manual liquid unloadings

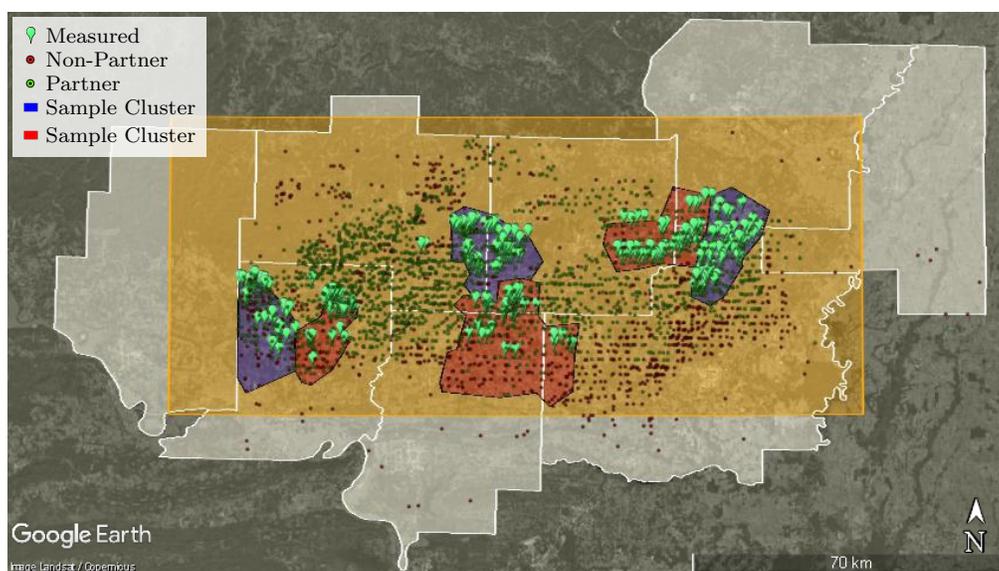


Figure 4.3: Modeled production sector emissions are based on the study on-site estimate (SOE) of Bell et al.⁷¹. Study area production facilities (well pads) were chosen for measurement using random sampling, in a clustered sampling strategy.

from measurements of horizontal wells without oil production classified as mid-continent in Allen et al.

Plunger Unloadings includes emissions from vented plunger unloadings, which may be triggered automatically or manually. Emissions from plunger unloadings are modeled using study partner provided activity data which included annual counts, and average plunger unloading durations. These activity data were spatially explicit and specific to individual wells. The GLAE utilizes emission rates for mid-continent plunger unloadings measured in Allen et al.⁷²

Fugitives as used in herein for the production sector, refers to the sum of Onsite Direct Measurements and Observed/Unmeasured sources as described in Bell et al.⁷¹

Pneumatics includes emissions from pneumatic devices present at production facilities based on study partner provided, spatially explicit, counts of pneumatic devices by type, per well. This category includes emissions from: pneumatic-powered chemical injection pumps; continuous high-bleed, continuous low-bleed, and intermittent-bleed pneumatic controllers.

Emission rates are simulated based measurement data from Allen et al.⁷³ for pneumatics in operation in the mid-continent region.

Compressors refers to combustion slip CH₄ emissions from compressor engines located at production facilities. Other compressor-related emissions are included in *Pneumatics*, or *Fugitives*, as applicable.

Table 4.2: Production sector source categories that contribute to modeled CH₄ emission rates predicted by the ground level area estimate (GLAE).

Production Model Categories
Manual Unloadings
Plunger Unloadings
Fugitives
Pneumatics
Compressors

4.3.2 GATHERING

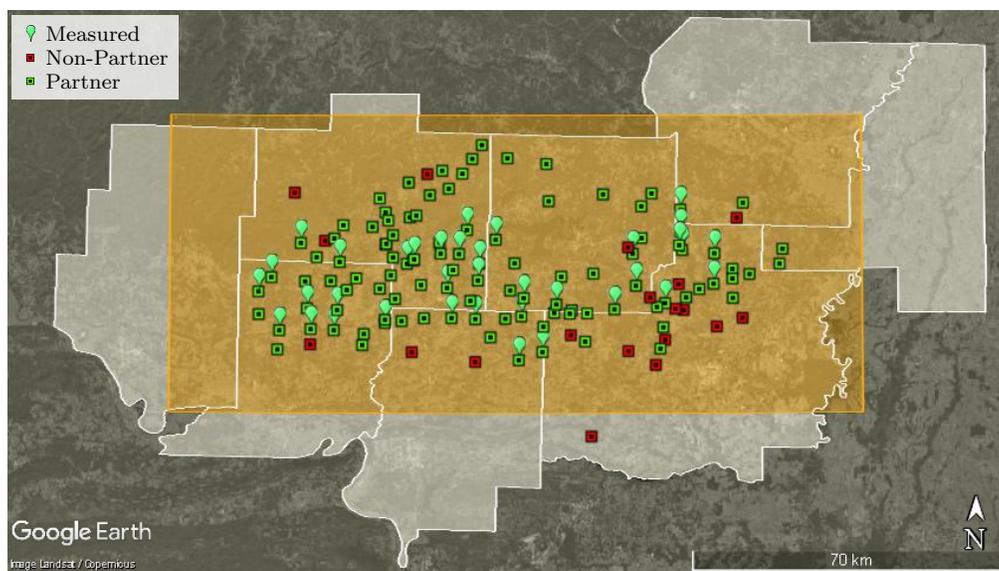


Figure 4.4: Modeled gathering sector emissions are based on the study on-site estimate (SOE) of Vaughn et al.⁷⁴. Study area gathering stations were chosen for measurement at random from facilities with suitable downwind road access for tracer flux measurements. Nearly all suitable facilities were measured.

Study partners own or manage 99 (79%) of the 125 gathering stations located within the study area and provided detailed activity data including facility locations, major equip-

ment inventories, and operating logs. Activity data for non-partner gathering stations were obtained from Arkansas Department of Environmental Quality (ADEQ) permit records; facility locations and compressor engine counts were confirmed using Google Earth. Methane emissions from gathering stations were estimated using on-site measurements, tracer measurements, aircraft measurements, and engineering estimates in a Monte Carlo model based on the SOE model described in Vaughn et al.⁷⁴ The SOE was extended to calculate emissions from unmeasured facilities, and a sub-model was added to capture emissions from non-routine sources. The non-routine emissions sub-model is based on tracer and aircraft measurements of atypical operating conditions (intended or unintended) made during this study. Emissions were calculated for source categories shown in Table 4.3 using the methods described in the following sections.

Table 4.3: Gathering sector source categories that contribute to modeled CH₄ emission rates predicted by the ground level area estimate (GLAE).

Gathering Model Categories
Component or Device Leaks and Losses
Combustion Slip
Crankcase Vents
Dehydrator Regenerator Vents
Compressor Engine Start-ups
Tank Venting
Gathering Lines

Component or Device Leaks and Losses (hereafter “leaks”) refer to ODMs and SDMs of sources as described in Vaughn et al.⁷⁴ ODMs refer to measurements made by on-site teams during the field campaign using high-flow samplers (Bacharach Hi Flow[®]). ODMs were made of dry gas sources spanning the measurable range of the high-flow sampler (0.05 SCFM–8 SCFM or equivalently 0.058–9.24 kg/h).⁴³ SDMs provide an emission rate estimate when ODMs were attempted but outside the measurable leak rate of the high-flow sampler, or when sources were observed with OGI but were not safe or accessible for measurement. Simulated direct measurements are re-sampled from ODMs of the same major equipment

category. Measured and unmeasured leaks observed with OGI and estimated to be within the measurable range of the high-flow sampler are termed “leak observations”.

Table 4.4: All on-site direct measurements made at gathering stations during the field campaign were assigned to one of the following categories.

Gathering Major Equipment Categories
Compressor
Dehydrator
Other
Pig Launcher/Receiver
Piping or Gas Line
Separator
Tank

At measured gathering stations CH₄ emissions from leaks were calculated as described in Vaughn et al.⁷⁴ To estimate leaks at un-measured gathering stations, leak count distributions were developed by dividing leak observation counts by major equipment counts at each measured facility. For example, all leak observations on dehydrators (excluding regenerator vents) at a measured facility were divided by the number of dehydrators at the facility, resulting in a distribution of dehydrator leaks per dehydrator. Leak observations from all other major equipment categories were normalized similarly using compressor engine counts. Compressor leaks were further disaggregated to distinguish rod packing vent and pressure relief valve emissions from other leaks.

For each Monte Carlo iteration, i , methane emissions from leaks at un-measured facility j are calculated in a two-step process. First, the number of leak observations is simulated for each major equipment category as:

$$N_{leakobs,i} = \sum_{k=1}^N \text{round}(\text{draw}(\text{Dist}) \cdot N) \tag{4.1}$$

Where:

N is the count of major equipment category k at facility j (compressors or dehydrators)

draw(Dist) indicates drawing one value at random from the distribution of normalized leak observations for major equipment category k

The result is a simulated leak observation count for each major equipment category at an unmeasured facility. The CH₄ emission rate from leaks in major equipment category k is then simulated based on the leak observation count as:

$$\dot{m}_{leaks,i} = \sum_{k=1}^{N_{leakobs}} \text{simulate}(leakobs_k) \quad (4.2)$$

Where:

$N_{leakobs}$ is the count of leak observations simulated for major equipment category k in equation 4.1

$\text{simulate}(leakobs_k)$ indicates simulating a leak observation as described in Vaughn et al.⁷⁴

Combustion Slip refers to unburned fuel entrained compressor engine exhaust. Combustion slip was not measured in this study; however, study partners provided engine exhaust stack test data for 111 engines located within the study area tested in the year prior to the field campaign. Tests were performed by measurement contractors using standard methods (EPA Method 19⁴⁵, EPA Method 320⁴⁶). Of the 111 engines tested, 24 were from one engine series (Caterpillar[®] G3500, rated at ≈ 1 MW), and 87 from another (Caterpillar[®] G3600, rated at ≈ 1.3 MW). Activity data from study partners and ADEQ indicate that the study area contains 447 gathering compressor engines, 416 of which belong to one of these two engine series. These tests therefore represent nearly one fourth of the compressor engines at gathering stations within the study area and nearly all (93%) compressor engines belong to one of these engine series, leading to high confidence in combustion slip estimates. All engines belonging to the two series tested were simulated using emission factors developed from test data. The 31 gathering compressor engines within the study area that did not

belong to one of these engine series were simulated using EPA AP-42⁷⁵ factors relevant to the engine classification.

Study partners also provided activity data for compressor engines that included run-hours, start-up times, and shut-down times for approximately 70% of gathering compressor engines within the study area. Combustion slip emissions are calculated for each hour of the study period using this activity data. Run hours and start-ups and shut-downs were applied directly to the engines they were provided for; all other engines were simulated by re-sampling from this data.

For each Monte Carlo iteration, i , combustion slip methane emissions for facility j were calculated as:

$$\dot{m}_{comb\ slip,i} = \sum_{k=1}^{N_{op}} EF_k \cdot \text{draw}(Load_k) \cdot RatedHP_k \quad (4.3)$$

Where:

N_{op} represents the count of compressor engines operating on-site for the hour simulated, whether known explicitly or simulated by re-sampling

EF_k is the emission factor relevant to engine k . EF_k is re-sampled from study partner provided test data for Caterpillar[®] G3500, and G3600 series engines. AP-42 factors are used otherwise.

$\text{draw}(Load_k)$ indicates drawing a fractional load at random from the distribution of operating loads observed during the field campaign, and applying it to engine k

$RatedHP_k$ is the rated power output of engine k

Crankcase Vents account for CH₄ vented from compressor engine crankcases because of imperfect piston ring sealing. Crankcase vents were simulated based on a Caterpillar[®] crankcase ventilation system application guide⁵¹; crankcase vents were not measured in this study. Expected crankcase vent hydrocarbon emissions are normally 3% of exhaust emissions at engine mid-life, but could reach 20% due to engine wear. Crankcase vent emissions

were simulated by multiplying combustion slip by a factor drawn at random from a normal distribution (mean 3%, assumed standard deviation 2%).

Dehydrator Regenerator Vents were simulated using the emission factor for dehydrators with flash tank vapor recovery from a 1996 GRI study⁵⁰ (0.003 (-52%/+102%) kg/h CH₄ per MMscf per day of gas processed). Most study partner dehydrators are equipped with flash tank vapor recovery, an emission control technique. The volume of gas processed is directly related to operating compressor engine horsepower, and was estimated on this basis.

For each Monte Carlo iteration, i , methane emissions from measured glycol dehydrator still vents at facility j are calculated as:

$$\dot{m}_{measdehy,i} = \begin{cases} \sum_{k=1}^N f_i \cdot ODM_{stillvent,k} & \text{if measured,} \\ 0 & \text{otherwise} \end{cases} \quad (4.4)$$

Where:

N is the number of on-site direct measurements of dehydrator still vents made at facility j not subject to any emission rate exceptions

f_i is a factor drawn from a normal distribution to account for the high-flow sampler measurement uncertainty ($\pm 10\%$)⁴³

Compressor Engine Start-ups account for emissions released from gas pneumatic starters and pumps used to start compressor engines. Study partners provided an estimate of 3800 scf of gas released per engine start. Emissions are simulated by drawing a value at random from a triangular distribution centered at 3800 scf, and ranging from 500 scf–5000 scf. Engine start-up times and locations are known for 70% of study area engines, and are simulated otherwise.

Tank Venting refers to continuous emissions from tanks in excess of the measurable leak rate of the high-flow sampler. This scenario was encountered on two occasions during the field campaign and both are simulated in the GLAE. In one instance, the aircraft team noted

significant CH₄ enhancement from a gathering station during a raster flight. The facility was measured³⁸ on three days (October 2st, 3rd, and 14th, 2015) with emission rates of 276 (\pm 99 kg/h), 676 (\pm 119 kg/h), and 739 (\pm 107 kg/h) on each day, respectively. Tracer and on-site measurements were made at this facility on October 6th, 2015. The source was identified as a produced water tank and the cause was identified as an open (hand-operated) valve on a compressor engine fuel scrubber. The tracer team was not able to make a complete facility measurement, but was able to measure the portion of the facility where the tank was located both with the valve open, and after it had been identified and closed. Subtracting the tracer estimates made in each operating state leads to an estimated 606 (\pm 278 kg/h) originating from the tank. On-site teams had no means to quantify or estimate an emission source of this magnitude.

For this facility only, on each Monte Carlo iteration, i , tank venting emissions are calculated by first randomly choosing a measurement day. If an aircraft measurement date is chosen, the emission rate for all other sources at the facility (as predicted by the SOE) is subtracted from the aircraft measurement and uncertainties are subtracted in quadrature. A random value is then selected from a triangular distribution centered at the difference, and bounded by the uncertainty (95% CI). If the tracer measurement date is chosen, a random value is selected from a triangular distribution described by the tracer measurement and associated uncertainty (95% CI). Tank emissions at this facility are a self-representing sample since the facility was not chosen for measurement randomly. Ground-based teams were dispatched to confirm the aircraft measurements. The aircraft did not identify any other facilities with persistent emission rates of this magnitude during the field campaign.

In another instance, tank venting was observed at a gathering station during random sampling. At this facility, the tracer team noticed significant CH₄ enhancement from a produced water tank, which on-site measurement teams confirmed as the source via OGI. The cause was not identified, but operators at the facility suspected a stuck dump valve. Tank venting emissions were estimated by subtracting the SOE from the tracer measure-

ment at this facility, since the SOE estimates all sources except the tank venting, and the tracer measurement captures all sources including the tank venting. The estimated tank venting emission rate was 140 (± 40 kg/h). The tracer team did not identify similar tank venting emissions at other measured gathering stations. We assume that the emission rate and observed frequency are representative of tank venting emissions from gathering stations within the study area. Each simulated gathering station (except the self-representing facility described previously) is assigned tank venting emissions at the probability observed, approximately 1 in 30. If a gathering facility is assigned tank venting emissions, an emission rate is drawn at random from a triangular distribution described by the estimated emission rate and associated uncertainty.

Gathering Pipelines herein refer to both underground pipelines and associated above ground equipment, and are simulated as described in Zimmerle et al.³³. During the field campaign, 96 kilometers of gathering pipelines and associated above ground equipment were surveyed and measured, including 56 pigging facilities and 39 block valves. Only one underground pipeline leak was identified and it accounted for 83% (4 kg/h) of measured emissions from gathering pipelines. Leaks were found most often on above-ground equipment. Zimmerle et al. estimate total study area CH₄ emissions from gathering pipelines of 400 kg/h (+214%/-87%, 95% CI).

For each Monte Carlo iteration, total methane emissions from gathering pipelines in the study area are calculated using the method described in Zimmerle et al.³³ Total emissions are then distributed to grid cells using a correlation based on the spatial density of wells. CH₄ emissions from gathering stations are assigned to the grid cells containing them.

4.3.3 TRANSMISSION

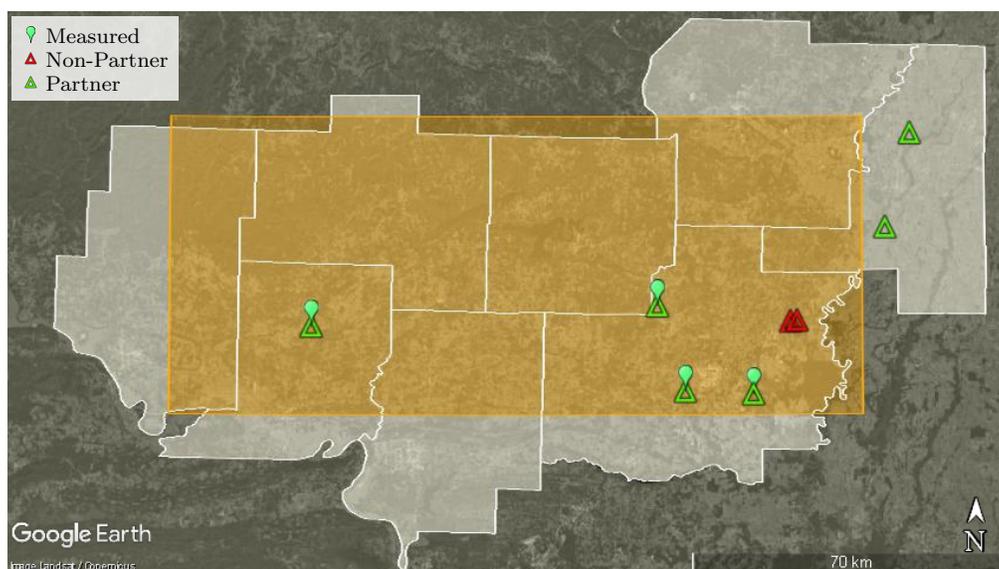


Figure 4.5: Study partners operated four of the six transmission stations located in the study area. Emissions from these four stations were modeled based on tracer measurements made during this study. Emissions from the two non-partner transmission stations were modeled based on EPA GHGRP data.

Four study partner transmission stations and two non-partner transmission stations were identified within the study area using study partner data, GHGRP data, and ADEQ records. Methane emissions from study partner transmission stations were estimated using tracer measurements made during this study. Emissions from non-partner transmission stations were calculated from data reported to the EPA GHGRP. First, CH_4 emissions for stationary combustion reported under 40 CFR 98.33.⁴⁹ (“Subpart C”) were recalculated using AP-42⁷⁵ emission factors. These results were then added to emissions reported under 40 CFR 98.230⁷⁶ (“Subpart W”) and normalized to provide an annual average hourly emission rate for the non-partner facilities. On-site measurements were not made at transmission stations in this study. A 95% confidence interval of $\pm 50\%$ is assumed for these emission rates.

For each Monte Carlo iteration, i , CH_4 emissions from transmission stations are calculated as follows:

$$\dot{m}_{trans,i} = \sum_{m=1}^4 \text{draw}(\dot{m}_{meas,m}) + \sum_{r=1}^2 \text{draw}(\dot{m}_{ghgrp,r}) \quad (4.5)$$

Where:

$\text{draw}(\dot{m}_{meas,m})$ indicates drawing one emission rate at random from a normal distribution described by tracer measurement and associated uncertainty at each of four measured facilities

$\text{draw}(\dot{m}_{ghgrp,r})$ indicates drawing one emission rate at random from a triangular distribution centered at the calculated annual average hourly emission rate, with assumed 95% CI of $\pm 50\%$

Calculated emissions are assigned spatially to the transmission category in the grid cells that contain the facilities.

4.3.4 DISTRIBUTION

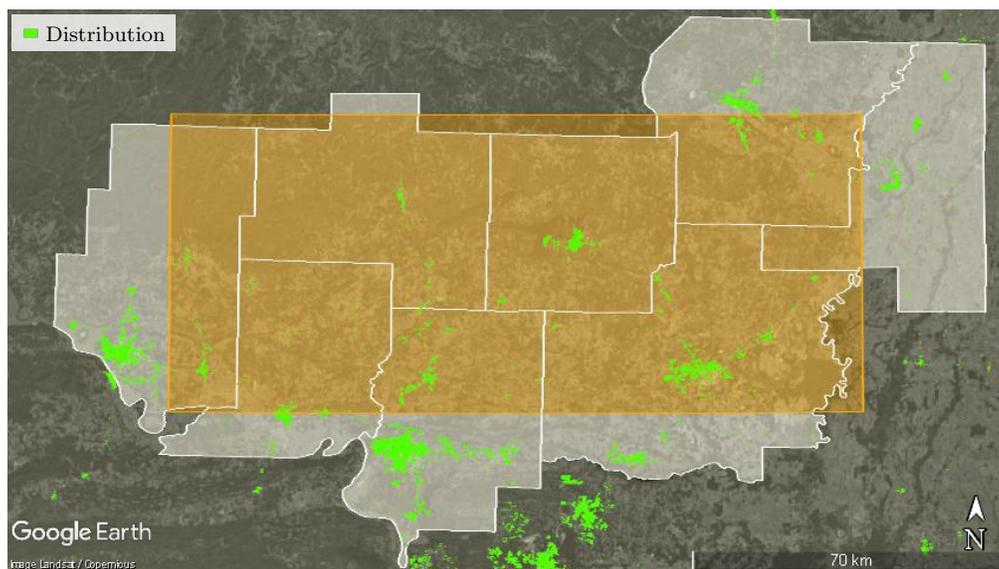


Figure 4.6: Distribution sector activities are concentrated in urban and suburban regions. One study partner distribution company served the entire study area.

Methane emissions from the distribution sector were estimated based on direct measurements performed during this study and activity data provided by study partners for most source categories. Sources with few or no measurements were estimated using activity data and emission factors from this and prior studies. One distribution company serves the entire

study area, enabling measurement across the entire industry sector. Distribution operations are concentrated mainly in urban and suburban areas with higher population density, as highlighted in Figure 4.6.

Leaks were measured at distribution facilities and on distribution pipelines. Distribution facilities are classified as transmission distribution transfer stations (TDTSs), metering and regulating (M&R) stations, or customer meters, while pipelines are classified as service mains, or service pipelines. Gas from transmission pipelines enters the TDTS on the “transmission side” and the pressure is reduced (from $\sim 1,000$ psi to $\sim 100\text{--}500$ psi) as the gas flows to the “distribution side” and enters the distribution system. A TDTS may contain equipment owned and operated by both the transmission and distribution operators, for example both operators typically measure gas flow during the custody exchange. Gas exiting the TDTS is routed to service mains which deliver it to M&R stations, where the gas flow is measured (“metering”) and pressure is further reduced. Metering was not performed at M&R stations within the study area because the system was wholly owned by a single operator. Gas exiting M&R stations is routed to service pipelines that deliver it to customer meters at commercial or residential locations.

Table 4.5: Distribution sector source categories that contribute to modeled CH_4 emission rates predicted by the ground level area estimate (GLAE).

Distribution Model Categories
Transmission Distribution Transfer Stations
Service Main Pipelines
Service Pipelines
Metering and Regulating
Commercial Sales Meters
Residential Sales Meters

Measured TDTS and M&R stations were grouped into three categories based on the gas pressure at the inlet to the facility. At some TDTSs the transmission side of the facility was not measured because study personnel did not have right-of-access at the start of the field campaign. Therefore, the transmission and distribution sides of TDTSs were modeled

independently to ensure inclusion of potential emissions at stations where the transmission side was not measured. Leak surveys were not performed to identify pipeline leaks. Pipeline leaks targeted for measurement were selected at random from a list of reported or identified leaks maintained by the partner company. This list was assumed to contain all distribution pipeline leaks within the study area that may have existed during the study period. A detailed description of the distribution measurements made during this study are provided by Pickering.⁷⁷ Only a small number of sales meters included in the reported leak list were measured during the study. Emission estimates for sales meters are therefore based on measurements made in this study, and a prior study which measured a large number of commercial and residential sales meters.⁷⁸

Table 4.6: Counts of measurements made at distribution facilities during the field campaign.

County	M&R ^a	Pipelines ^b		TDTS ^a	
		Mains	Services	Distribution Side	Transmission Side
Cleburne	5/5	0/1	0/0	6/6	6/6
Conway	10/10	0/0	0/0	7/8	6/8
Faulkner	30/37	2/3	5/11	9/11	9/11
Independence	0/47	0/5	0/6	0/3	0/3
Jackson	0/27	0/5	0/1	0/2	0/2
Pope	15/15	1/4	5/9	4/5	4/5
Van Buren	27/27	0/0	0/0	1/1	0/1
White	13/29	11/23	10/17	2/6	0/6

^a Measured facilities / total facilities

^b Measured leaks / reported or otherwise identified leaks

For each Monte Carlo iteration, i , CH_4 emissions from distribution facility category k in county j are calculated as follows:

$$\dot{m}_{category(k),i} = (\text{draw}(EF_k) \cdot AD_k + MEAS_{category(k)}) \cdot (Area \cap_j) \quad (4.6)$$

Where:

$\text{draw}(EF_k)$ indicates drawing one emission rate at random for facility $category(k)$

AD_k indicates the activity data (facility count) for $category(k)$ for county j

$MEAS_{category(k)}$ is the sum of all measurements for $category(k)$ in county j

$Area \cap_j$ is the fraction of county j that spatially intersects the study area

Emissions are assumed to be distributed uniformly throughout the regions with distribution service (Figure 4.6) and are apportioned to grid cells in a two-step process. First, emissions from distribution operations in county j are scaled by the fractional area of distribution operations in county j that spatially intersect the study area. Second, emissions are apportioned to grid cells by the fractional area of distribution operations that spatially intersect an individual grid cell. In this way, county level activity data, and measured and simulated emissions are concentrated in regions with distribution operations, and are scaled by the overlap with the study area. This also allows emissions to be attributed appropriately to grid cells that intersect distribution operations in multiple counties.

4.4 NON-OIL AND GAS RELATED METHANE SOURCES

4.4.1 LIVESTOCK

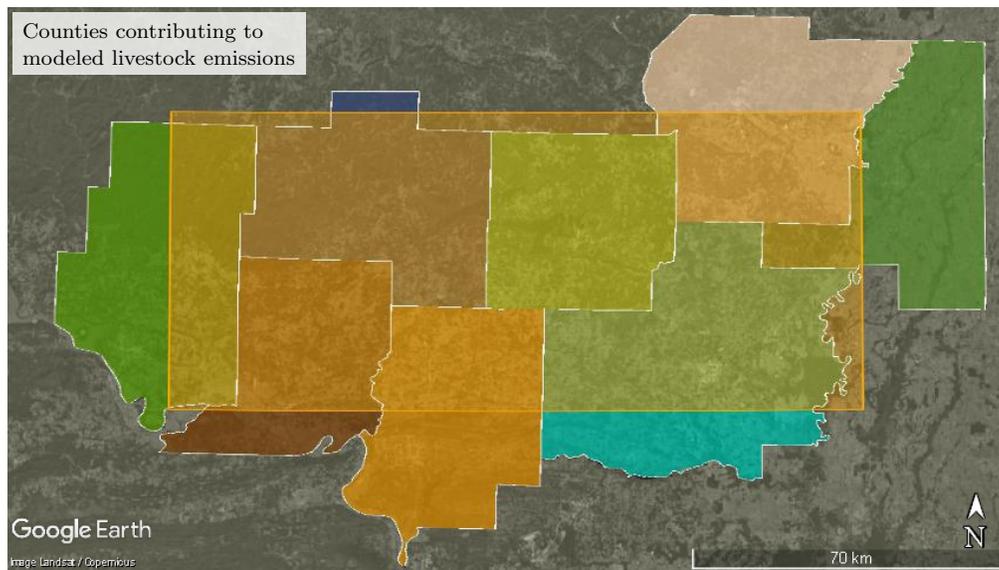


Figure 4.7: Livestock data were only available at the county-level, and were apportioned to the study area (orange rectangle) in proportion to spatial intersection with the counties shown.

Methane emissions from livestock were calculated using activity data from the United States Department of Agriculture (USDA) census and emission factors from the U.S. EPA GHGI⁸, and Intergovernmental Panel on Climate Change (IPCC)⁷⁹ guidelines. Livestock counts were obtained at the county level from the 2012 USDA census⁸⁰ for the eight Arkansas counties that significantly overlap the study area. Data were not available for all source categories for all counties because data is withheld in cases where it can be attributed to a unique producer. In cases where 2012 data were withheld, 2007 data were used instead. If neither 2012 nor 2007 data were available for a category, its activity data was considered 0 in this model.

Table 4.7: 2012 USDA livestock census data for study area counties.

USDA County-Level Activity Data							
County	Beef Cows	Milk Cows	Other Cattle	Hogs	Layers	Pullets	Broilers
Cleburne	13 606	0	206 706	140	389 627 ¹	134 030 ¹	1 991 264 ¹
Conway	20 303 ¹	1130 ¹	29 718 ¹	12 512	62 928	114	6 888 751
Faulkner	14 390	886	14 892	129	2525	196	76
Independence	19 533	0	16 520	104	367 690	718 857	6 665 939
Jackson	2288	0	2170	0	386	D ²	D ²
Pope	16 181	0	13 689	9380	155 763	303 221	4 871 203
Van Buren	11 135 ¹	790 ¹	8372 ¹	3103 ¹	1031	164	489 312
White	20 234	401	21 316	408	D ²	D ²	806 465

¹ 2007 USDA census data used

² Withheld to avoid disclosing data for individual farms

The USDA census inventories cattle as ‘beef cows’, ‘milk cows’ and ‘other cattle’. Emission factors are available from the GHGI⁸ for ‘dairy cattle’ and ‘beef cattle’. For this reason, ‘other cows’ from the AR USDA county level census data were redistributed proportionally to the ‘milk cow’ and ‘beef cow’ categories. The only poultry considered in this model were chicken. Chicken are inventoried in the USDA census as ‘layers’, ‘pullets’ and ‘broilers’. Pullets grow to be layer flock replacements and were therefore added to the layer inventory in this model. No uncertainty was applied to livestock activity data.

Emission factors used for livestock categories considered in the model are shown in Table 4.9. Emission factors are the U.S. implied emission factors developed in the GHGI⁸, and uncertainties are 95% confidence intervals provided in the IPCC⁷⁹ guidelines for GHGIs.

Table 4.8: 2012 USDA livestock census data for study area counties, as modeled.

Modeled County-Level Activity Data					
County	Beef Cows	Milk Cows	Hogs	Layers	Broilers
Cleburne	34 276	0	140	523 657	1 991 264
Conway	48 454	2697	12 512	63 042	6 888 751
Faulkner	28 418	1750	129	2721	76
Independence	36 053	0	104	1 086 547	6 665 939
Jackson	4458	0	0	386	0
Pope	29 870	0	9380	458 984	4 871 203
Van Buren	18 952	1345	3103	1195	489 312
White	41 136	815	408	0	806 465

Table 4.9: Emission factors and uncertainty used in the model for estimating CH₄ emissions from livestock.

Livestock Emission Factors Used In Model		
Category	CH ₄ Emission Factor (g/head/hr) ^a	95% Confidence Interval ^b
Beef Cattle Enteric Fermentation	8.4	±50%
Beef Cattle Manure Management	0.2	±30%
Dairy Cattle Enteric Fermentation	13.5	±50%
Dairy Cattle Manure Management	8.0	±30%
Swine Enteric Fermentation	0.2	±50%
Swine Manure Management	1.6	±30%
Poultry Manure Management	0.01	±30%

^a US EPA GHGI⁸^b IPCC guidelines⁷⁹

For each Monte Carlo iteration, i , CH₄ emissions from livestock category k in county j are calculated as follows:

$$\dot{m}_{category(k),i} = \text{draw}(EF_k) \cdot (AD_k) \cdot (Area \cap_j) \quad (4.7)$$

Where:

$\text{draw}(EF_k)$ indicates drawing one emission factor value at random from a triangular distribution centered at EF_k , and bounded by its associated confidence interval, as shown in Table 4.9

AD_k indicates the activity data (head count) for $category(k)$ for county j

$Area \cap_j$ is the fraction of county j that spatially intersects the study area

Emissions are assumed to be distributed uniformly throughout the county area and are apportioned to grid cells in a two-step process. First, emissions from county j are scaled by the fractional area of county j that spatially intersects the study area, resulting in a sub-county emission estimate. Second, sub-county emissions are apportioned to grid cells by the fractional area of sub-county j that spatially intersects an individual grid cell. In this way, county level emissions are scaled by the area overlap with the study area and emissions are attributed appropriately to grid cells that intersect multiple counties.

4.4.2 RICE CULTIVATION

Methane emissions from rice cultivation were calculated based on a combination of IPCC factors⁷⁹ and USDA county level census data for the state of Arkansas.⁸⁰ Arkansas has the largest area of rice harvested in all U.S. states.⁸ However, the majority of CH₄ emissions from rice cultivation occur during the growing season when fields are flooded. Rice is typically harvested in early September, and was thus likely harvested before the mass balance flights which occurred on October 1st and 2nd. One study of Arkansas rice fields⁸¹ found that post-harvest CH₄ emissions represented 2% of annual emissions. Therefore we have multiplied the IPCC rice emission factor by 0.02 to develop a study relevant CH₄ emission factor for rice cultivation.

Table 4.10: The emission factor for rice cultivation used in the GLAE is based on IPCC guidelines, modified to represent post-harvest CH₄ emissions.

Rice Cultivation Emission Factor Used In Model		
	CH ₄ Emission Factor	
Emission Source	(kg/hr/m ²)	95% Confidence Interval
Rice Cultivation ^a	108×10^{-9}	-39%/ + 70%

^a IPCC⁷⁹ default emissions factor modified by Smartt et al.⁸¹

For each Monte Carlo iteration, i , CH_4 emissions from rice cultivation in county j are calculated as follows:

$$\dot{m}_{rice,i} = \text{draw}(EF_{rice}) \cdot (AD_{rice}) \cdot (Area \cap_j) \quad (4.8)$$

Where:

$\text{draw}(EF_{rice})$ indicates drawing one emission factor value at random from a triangular distribution centered at EF_{rice} , and bounded by its associated confidence interval

AD_{rice} indicates the activity data (area harvested) for county j

$Area \cap_j$ is the fraction of county j that spatially intersects the study area

Emissions are assumed to be distributed uniformly throughout the county area and are apportioned to grid cells in a two-step process. First, emissions from county j are scaled by the fractional area of county j that spatially intersects the study area, resulting in a sub-county emission estimate. Second, sub-county emissions are apportioned to grid cells by the fractional area of sub-county j that spatially intersects an individual grid cell. In this way, county level emissions are scaled by the area overlap with the study area and emissions are attributed appropriately to grid cells that intersect multiple counties.

4.4.3 WETLANDS

Methane emissions from wetlands are calculated based on activity data from the U.S. Fish & Wildlife Service⁸² and emission rates from a variety of sources. Geospatial data for land area containing permanently flooded emergent and forested wetlands, lakes, ponds and rivers were extracted from shapefiles downloaded from the national wetlands inventory.⁸² Temporarily and seasonally flooded areas were not considered because the mass balance flights occurred during the dry season, on clear days during a period of little rainfall.

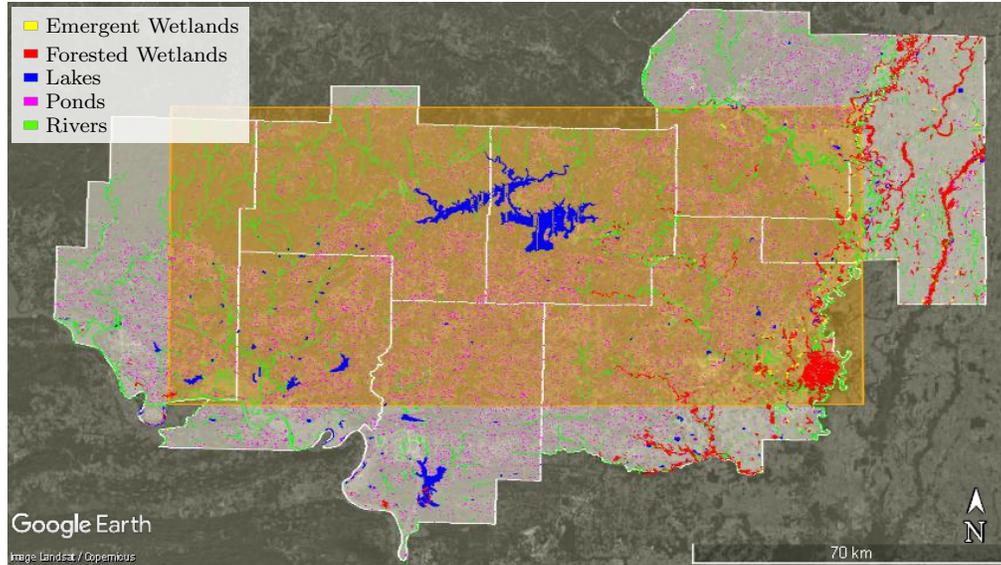


Figure 4.8: Wetlands considered within the study area.

Table 4.11: Central, lower and upper bounds for triangular distributions used in wetland emission factor simulations.

Wetland Emission Rates Used In Model			
Category	Central Estimate (kg/hr/m ²)	Lower Bound (kg/hr/m ²)	Upper Bound (kg/hr/m ²)
Forested Wetlands	3.75×10^{-6}	1.7×10^{-6}	6.7×10^{-6}
Emergent Wetlands	6.7×10^{-6}	4.25×10^{-6}	10.8×10^{-6}
Lakes	1.04×10^{-6}	1.0×10^{-6}	4.7×10^{-6}
Ponds	0.76×10^{-6}	0.4×10^{-6}	1.1×10^{-6}
Rivers	0.55×10^{-6}	-2.8×10^{-6}	3.9×10^{-6}

A range of emission rates for temperate and subtropical forested and emergent wetlands were obtained from Bartlett et al.⁸³ Deemer et al.⁸⁴ show that CH₄ emission rates are correlated with chlorophyll *a* concentrations. Chlorophyll *a* concentration measurements for Greers Ferry lake, the largest within the study area, were obtained from the Arkansas Department of Environmental Quality (ADEQ)⁸⁵. A central estimate for CH₄ emission rates from lakes within the study area was made by comparing the chlorophyll *a* concentrations in Greers Ferry lake with the range of CH₄ concentrations and fluxes in Deemer et al., as described in Pickering.⁷⁷

A recent study by Holgerson and Raymond⁸⁶ found that CH₄ fluxes from small ponds increased with decreasing surface area. They provide CH₄ flux rates for lakes and ponds of varying size class. The central estimate used in the model is a weighted average of these flux rates and the size class of all ponds within the study area. The lower and upper bounds are a weighted average of their reported standard error, expanded to two sigma.

Methane emissions rates for rivers in the study area are based on total CH₄ emissions, and total surface area for rivers located between 25°–54° latitude provided in Bastviken et al.⁸⁷ Lower and upper bounds were estimated by expanding their stated uncertainty on total CH₄ emissions to two sigma.

For each Monte Carlo iteration, i , CH₄ emissions from wetland category k are calculated as follows:

$$\dot{m}_{wetland(k),i} = \text{draw}(EF_k) \cdot (AD_k) \quad (4.9)$$

Where:

$\text{draw}(EF_k)$ indicates drawing one emission factor value at random from a triangular distribution centered at EF_k , with associated lower and upper bounds as shown in Table 4.11

AD_k indicates the activity data (surface area) for grid cell m within the study area

Emissions for each wetland category are assumed to be distributed uniformly throughout each containing grid cell. Total surface area for each wetland category within each grid cell is calculated directly by spatial intersection. No intermediate allocation from county level to study area is required as was for livestock and rice cultivation.

4.4.4 GEOLOGIC SEEPS

Methane emissions from geologic seeps area calculated based on microseepage rates observed by Etiope et al.⁸⁸ Microseepage refers to positive CH₄ flux at the ground surface due to gas migration from underground gas reservoirs, which can potentially occur in sedimentary basins in dry climates with underlying gas or petroleum reservoirs.⁸⁸ Microseepage emission rates are categorized in three levels by Etiope et al. Level 1 seepage exceeds 50 mg/m²/day, level 2 seepage ranges from 5–50 mg/m²/day, and level 3 seepage ranges from 0–5 mg/m²/day. In this study, the mean, lower and upper bounds for level 3 seepage were applied to the study area.

Table 4.12: Central, lower and upper bounds for triangular distributions used in geologic seep emission factor simulations.

Geologic Seep Emission Rates Used In Model			
Category	Central Estimate (kg/hr/m ²)	Lower Bound (kg/hr/m ²)	Upper Bound (kg/hr/m ²)
Microseepage ^a	58×10^{-9}	0	208×10^{-9}

^a Corresponds to level 3 seepage in Etiope et al.⁸⁸

For each Monte Carlo iteration, i , CH₄ emissions from geologic seeps are calculated as follows:

$$\dot{m}_{seep,i} = \text{draw}(EF_{seep}) \cdot (AD_{seep}) \quad (4.10)$$

Where:

$\text{draw}(EF_{seep})$ indicates drawing one emission factor value at random from a triangular distribution centered at EF_{seep} , with associated lower and upper bounds as shown in Table 4.12

AD_{seep} indicates the activity data (surface area) for grid cell m within the study area

Calculated geologic seep emission are apportioned uniformly to study area grid cells.

4.4.5 LANDFILLS

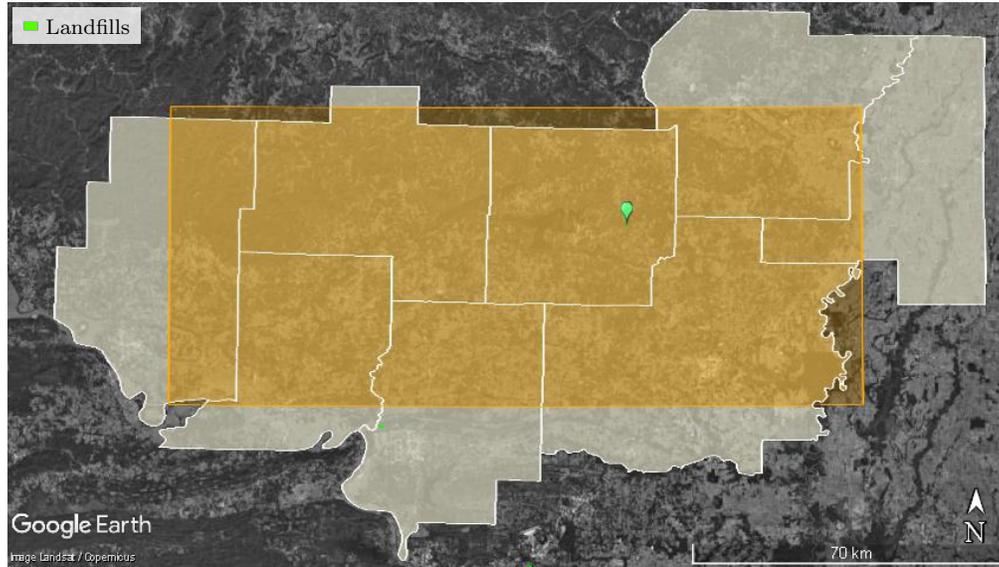


Figure 4.9: Only one landfill (green balloon) was identified within the study area (orange highlighting).

Methane emissions from landfills are based on six measurements of landfills made by the aircraft during the field campaign, one of which was measured twice. The five measured landfills were not within the study area boundary, but were reported to the GHGRP. Measured rates and 95% confidence intervals are shown in Table 4.13, along with hourly rates calculated from annual CH_4 emission reported to GHGRP. Landfill areas were estimated using Google Earth, and emission factors were created based on the rate measured by the aircraft, and the estimated area. The study area only contained one landfill to the authors' knowledge, and this landfill was not reported to the GHGRP. The area of the landfill was also estimated using Google Earth, and the developed emission factors were applied in the Monte Carlo Model as follows.

For each Monte Carlo iteration, i , CH_4 emissions from the landfill are calculated as follows:

$$\dot{m}_{\text{landfill},i} = \text{draw}(EF_{\text{landfill}}) \cdot (AD_{\text{landfill}}) \quad (4.11)$$

Where:

Table 4.13: Landfill emission factors measured by aircraft compared to GHGRP reported average rates. Areas were estimated using Google Earth.

Landfill	Date	Aircraft Estimate (kg/h)	GHGRP (kg/h)	Area (m ²)	Study EF (kg/h/m ²)
Conway	9/25/2015	251.1±59.6	627.0	444 920	5.64×10^{-4}
Conway	10/13/2015	263.9±37.8	627.0	444 920	5.93×10^{-4}
Little Rock City	10/13/2015	1105.6±141.6	172.2	649 973	1.70×10^{-3}
Modelfill	10/13/2015	18±2.3	35.2	558 079	3.23×10^{-5}
Two Pine	10/13/2015	788±177	277.5	1 168 453	6.74×10^{-4}
Saline	10/13/2015	441.9±107	627.0	437 173	1.01×10^{-3}

$\text{draw}(EF_{\text{landfill}})$ indicates drawing one emission rate at random from the six landfill measurements made in the study. Uncertainty is then considered drawing a new emission rate from a triangular distribution centered at the measured emission rate drawn, and bounded by its associated confidence interval, as shown in Table 4.13. This emission rate is then normalized by the estimated area of the measured landfill resulting in EF_{landfill}

AD_{landfill} indicates the activity data (surface area) for the simulated landfill

Calculated emissions are then assigned to the landfill category in the grid cell that contains the landfill.

4.4.6 WASTEWATER TREATMENT

Methane emissions from wastewater treatment are based on 2015 population estimates for study area counties from the U.S. Census⁸⁹, and septic usage estimates from the National Environmental Services Center⁹⁰. Sewer and septic use are provided on a per household basis and we have assumed an equivalent ratio on a per person basis.

Emission factors are developed from a study on residential septic systems by Leverenz et al.⁹¹, and from the GHGI⁸ for centralized sewer systems. The GHGI estimates that 80% of the U.S. population is served by centralized sewer systems. Total CH₄ emissions from sewer system were divided by 80% of the U.S. population resulting in an emission factor of 1.3 g

Table 4.14: Wastewater activity data used in the model.

Modeled County-Level Activity Data			
County	Population	Households with Central Sewer (%)	Households with Septic Systems (%)
Cleburne	25 467	29	68
Conway	21 019	40	58
Faulkner	121 552	49	50
Independence	12 898	35	64
Jackson	17 338	63	35
Pope	63 390	51	48
Van Buren	16 771	25	70
White	79 161	51	48

* A portion of households in each county are served by other means

CH₄/day/person. Uncertainty was assumed to be the same as that provided for residential wastewater treatment, -37% / +8%.

Table 4.15: Wastewater emission factors used in the model.

Wastewater Emission Rates Used In Model			
Category	Central Estimate (g/day/person)	Lower Bound (g/day/person)	Upper Bound (g/day/person)
Septic Tanks ^a	11.0	6.3	17.9
Sewer Systems ^b	1.3	0.8	1.4

^a Central estimate is geometric mean of all sampled septic tanks. Upper and lower bounds are the geometric means of multiple measurements of individual tanks.⁹¹

^b Estimated from US Census and GHGI

For each Monte Carlo iteration, i , CH₄ emissions from wastewater treatment in county j are calculated as follows:

$$\dot{m}_{wastewater,i} = \text{draw}(EF_{sewer}) \cdot (AD_{sewer}) + \text{draw}(EF_{septic}) \cdot (AD_{septic}) \quad (4.12)$$

Where:

$\text{draw}(EF_{sewer})$ or $\text{draw}(EF_{septic})$ indicates drawing one emission factor value at random from a triangular distribution centered at EF_{sewer} or EF_{septic} , and bounded by its associated confidence interval

AD_{sewer} or AD_{septic} indicates the activity data (sewer or septic users) for county j

Emissions are assumed to be distributed uniformly throughout the county area and are apportioned to grid cells in a two-step process. First, emissions from county j are scaled by the fractional area of county j that spatially intersects the study area, resulting in a sub-county emission estimate. Second, sub-county emissions are apportioned to grid cells by the fractional area of sub-county j that spatially intersects an individual grid cell. In this way, county level emissions are scaled by the area overlap with the study area and emissions are attributed appropriately to grid cells that intersect multiple counties.

4.4.7 GHGRP FACILITIES

Facilities reporting to the EPA GHGRP in categories other than Petroleum and Natural Gas Systems were identified using the EPA Facility Level Information on GreenHouse gases Tool (FLIGHT)⁹². Only one facility within the study area was identified which was not accounted for in other categories within the model.

Table 4.16: GHGRP Facility emission factors used in the model.

GHGRP Facility Emission Rates Used In Model		
Facility	Reported Methane Emissions (tonne CH ₄ /yr in CO ₂ e)	Modeled Methane Emissions (kg CH ₄ /hr) ^a
Independence Power Plant	14 662	66.9

^a AR4 GWPs

For each Monte Carlo iteration, i , CH₄ emissions from GHGRP facilities, are calculated from reported methane emissions in tonne CH₄/yr CO₂_e, assuming IPCC fourth assessment report global warming potentials (GWPs), and 8,760 hrs. Emissions are assumed constant, and no uncertainty is applied. Emissions are assigned spatially to the GHGRP category in the grid cells that contain the facilities.

4.5 RESULTS AND DISCUSSION

4.5.1 TEMPORALLY VARYING GROUND-LEVEL AREA ESTIMATE (GLAE)

Bottom-up methane emissions from the study area were calculated in a Monte Carlo model on an hourly basis for the 48-hour period spanning October 1st and 2nd, 2015 as shown in Figure 4.10, and Appendix C.3. Each hourly result is the sum of a spatially resolved GLAE that considers emissions from thermogenic (oil and gas and non-oil and gas sources) and biogenic sources within the study area. Production, gathering, transmission and distribution sectors account for all oil and gas activities in the study area; there are no processing or storage facilities.

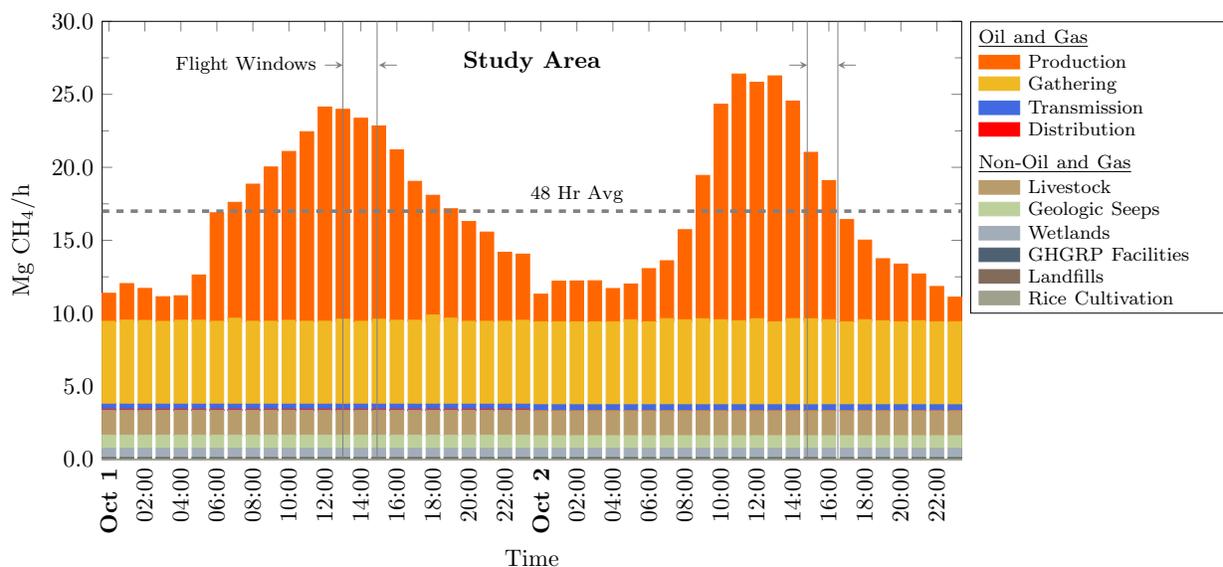


Figure 4.10: Hourly ground-level area estimates (GLAEs) predicted for days corresponding to aircraft mass balance flights show relatively steady emissions from most source categories. However, production emissions show significant diurnal variation due to manual liquid unloadings triggered by workers during the workday.

Emission rates from production facilities are based on the SOE of Bell et al.⁷¹, a comprehensive facility-level emission rate estimate. The production sector SOE uses emission factors developed from direct measurements made during this and other recent studies. Collectively, study partner companies operate 82% of active wells within the study area and

provided extensive activity data including equipment inventories and records of episodic emissions that occurred during the field campaign.

Results from most source categories in the GLAE model vary little in time; however, there is significant diurnal variation in emissions from the production sector. This variation is driven by MLUs initiated by workers in normal operations in the study area. “Unloading” a well refers to the variety of techniques used to remove accumulated liquids from the wellbore that slow gas production. There are many unloading techniques, some of which are not vented to atmosphere and do not result in emissions. Here, we consider MLUs, and plunger lift unloadings that vent to atmosphere. In a typical MLU a well is “shut in” to stop production and allow pressure to build downhole. When the well is brought back on-line, flow is diverted from pressurized separators to produced water tanks at atmospheric pressure. This practice reduces back pressure and increases flow velocity in the wellbore which aids in liquid removal. When the wellbore is cleared, the flow is reverted to separators and unloading emissions cease.

Emissions from MLUs alone increase the GLAE emission rate by a factor of two at the mid-day peak relative to night-time emissions, which only include emissions from MLUs in rare instances. Methane emission from MLUs were modeled based on data from Allen et al.⁷², which show that emission rates of manually triggered liquid unloadings vary greatly in time, with no consistent pattern. The time average emission rate varies by a factor of five (247—1253 kg/h, mean = 513 kg/h) among the mid-continent MLUs included in the GLAE model. While no direct measurement of MLUs were made in this study, tracer measurements were made at one production facility during an MLU and reported a methane emission rate of 810 (603—1018 kg/h 95%CI), which supports the use of the Allen et al.⁷² data. Activity data used to model MLUs was provided by study partners and included the location, start time, and duration of each individual MLU. Regardless of emission rate, activity data indicate an inevitable daily peak in emissions from MLUs; the shape of the daily peak is dependent on

the aggregate minute-by-minute emission rate of all active MLUs, in combination with other sources.

Emissions from gathering stations are steady and do not exhibit the diurnal pattern seen in the production sector. Gathering stations operate at high utilization rates and un-burned fuel entrained in compressor engine exhaust (“combustion slip”) is the largest emission source at gathering stations in aggregate. Emissions from gathering stations are divided into two sub-models: routine and non-routine emissions. Routine emissions from gathering stations are based on the SOE of Vaughn et al.⁷⁴ which utilizes on-site, tracer, and aircraft measurements made at 36 gathering stations during the field campaign. In addition, study partners operate 99 of the 125 gathering stations in the study area, and provided extensive activity data including facility locations, major equipment inventories, and activity data that included compressor engine start and stop times. Study partners also provided compressor engine exhaust test data from 111 compressor engines tested in the year prior to the study. All of these data are used to model routine emissions from gathering stations.

Non-routine emissions were included in the model based on observations made by tracer and aircraft measurement teams. Observations during this study indicate that abnormal process conditions (e.g. stuck dump valves, intentional or un-intentional venting) at any one station can exceed all other sources at the station combined. In general, the frequency of occurrence, size, and duration of these emissions remain an area of limited knowledge. Two sources of non-routine emissions are included in the model. First, in one case, a large (150—600 kg/h) source at a gathering station was identified repeatedly by the aircraft, on different days, during grid-patterned “raster” flights over the study area. Only one source was detected in this way, and given the thorough spatial coverage of the flights we assume it was the only source detectable in this way within the study area during the study period. Second, we assume that “significant tank venting” (e.g. stuck dump valve) emissions occurred at the

frequency observed by tracer teams, approximately 1 out of 30 stations at any one time, with similar emission rates.

Emissions from the transmission and distribution sectors contribute few emissions to the GLAE total. Four of the six transmission stations identified within the study area are operated by study partners and were modeled based on measurements made by tracer teams. The remaining two stations were modeled based on annual average emissions from GHGRP reported data⁹², where Subpart C⁴⁹ combustion emissions were replaced using appropriate EPA AP-42⁷⁵ factors. Distribution emissions are based on direct measurements made in this study, and activity data provided by study partners for most source categories. Leaks were measured at transmission distribution transfer stations, metering and regulating stations, customer meters, and on distribution pipelines (see Section 4.3.4). Distribution source categories with few or no measurements were estimated using activity data and emission factors from prior studies.⁷⁸

Finally, non-oil and gas emissions are based on a variety of sources and are calculated using an approach similar to Schwietzke et al.²³ (see Section 4.2), and are small relative natural gas production and gathering emissions, as shown in Figure 4.10.

Hourly GLAE simulation results were averaged for the 48-hour period spanning October 1st and 2nd, and the three-hour mid-day period (11–13 CDT) on each day. Resulting emission rates are reported for each source category in Table 4.17, along with total study area methane emission rates based on annual data from the GHGRP and GHGI (for the production and gathering sectors). For the production sector, GHGI emissions are calculated (excluding gathering) from production weighted national emissions, Arkansas well counts from the Arkansas Oil and Gas Commission (AOGC)⁹³, and study partner provided well count data. GHGRP production emissions are calculated from operators reporting to the “345-Arkoma” basin with operations in study area counties. Reported GHGRP production emissions are first scaled by GHGRP reported well counts within the study area, and then emissions-weighted by operator. This weighted, per-well, emission rate is then multiplied by

the same well count used in the GLAE model. GHGI gathering emissions are calculated by dividing net gathering emissions in the 2016 inventory⁸ by the facility count. This national, per-facility average is then multiplied by the number of gathering stations included in the GLAE model. GHGRP gathering emission are the sum of reported gathering stations in the study area for the 2015 reporting year.⁹⁴ Note that most gathering stations report under Subpart C of the GHGRP as direct emitters for stationary combustion (e.g. compressor engine) CO₂ emissions. The methane emission factor used in Subpart C reporting is known to be unrepresentative of actual combustion methane emissions from reciprocating compressor engines at transmission¹⁸ and gathering stations¹⁴.

Both the GHGI and GHGRP estimate lower emissions than the 48-hour GLAE modeled average (using the GLAE 48-hour average for categories other than gathering and production). Both inventory estimates are lower than GLAE modeled mid-day emissions by a factor of two–three.

Table 4.17: Ground-level area estimates compared to national inventories for the 48-hr period corresponding to the dates of aircraft mass balance flights, and the mid-day hours on each day.

Mg/h CH ₄	GLAE 48-hr Average (%)	Oct 1 GLAE 11:00–13:00 (%)	Oct 2 GLAE 11:00–13:00 (%)	GHGI	GHGRP
NG Production	7.4 (43.6)	14.0 (59.4)	16.6 (63.5)	4.9	6.8
NG Gathering	5.8 (33.9)	5.7 (24.3)	5.8 (22.0)	5.5	0.0
Livestock	1.7 (10.1)	1.7 (7.3)	1.7 (6.5)	1.7	1.7
Geologic Seeps	0.9 (5.2)	0.9 (3.8)	0.9 (3.3)	0.9	0.9
Wetlands	0.6 (3.8)	0.6 (2.7)	0.6 (2.4)	0.6	0.6
NG Transmission	0.4 (2.3)	0.4 (1.7)	0.4 (1.5)	0.4	0.4
Other GHGRP Facilities	0.1 (0.4)	0.1 (0.3)	0.1 (0.3)	0.1	0.1
Waste Water Treatment	0.1 (0.4)	0.1 (0.3)	0.1 (0.2)	0.1	0.1
Landfills	0.0 (0.2)	0.0 (0.2)	0.0 (0.1)	0.0	0.0
Rice	0.0 (0.1)	0.0 (0.0)	0.0 (0.0)	0.0	0.0
NG Distribution	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0	0.0
Study Area Total (95% CI)	17.0 (14.7-19.4)	23.6 (20.3-27.0)	26.2 (22.7-30.0)	14.2	10.7

4.5.2 COMPARISONS CONSIDERING FLIGHT TIME

Bottom-up methane emissions were also calculated for the times-of-day corresponding to top-down aircraft mass balance flights and confidence intervals overlap on both days, indicating agreement between the two estimates within statistical uncertainties, as shown in Figure 4.11, and Appendix C.1. Emissions aggregated by eastern and western portions of the study area also agree with the eastern and western results presented in Schwietzke et al.²³ Thus, bottom-up to top-down agreement is substantially improved relative to using disaggregated annual averages to model bottom-up emissions. The use of representative emission factors coupled with improved modeling of the timing and counts in activity data more accurately capture the emissions from events occurring during aircraft mass balance flights.

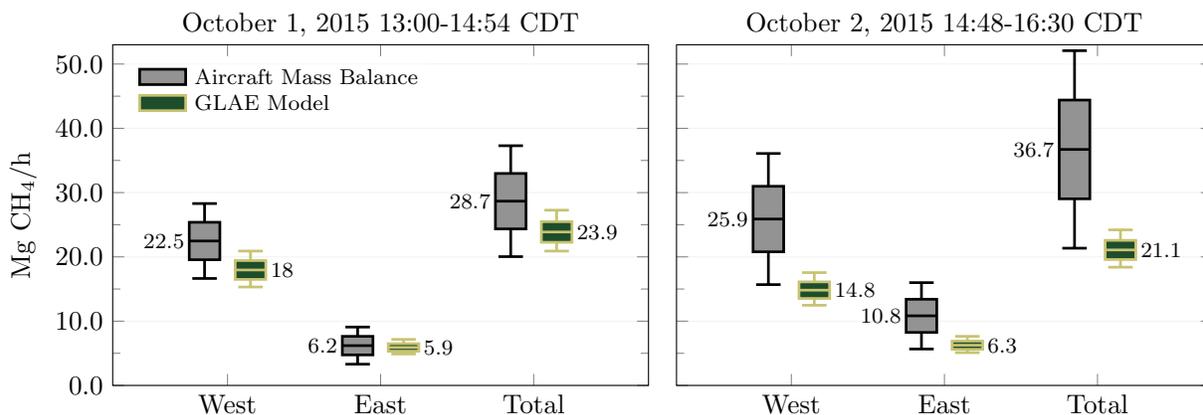


Figure 4.11: When time-aligned, ground-level area estimates produce overlapping confidence intervals with aircraft mass balance flight results, indicating agreement within statistical uncertainties. Top-down and bottom-up estimates also agree in the eastern and western sub-portions of the study area.

Emissions from each source category in the GLAE model were aggregated to 0.04° longitude (~ 3.8 km) grid cells, as shown by the orange squares in Figure 4.1. The total emission rate from each grid cell was idealized as a point source located in the center of each grid cell (white dots in Figure 4.12). Gaussian dispersion⁹⁵ was used to propagate emissions down-wind based on daily average wind speed and direction, as illustrated in Figure 4.12.

Contributions from each grid cell to each 0.02° longitudinal “bin” were summed by linear superposition, and an emission rate was calculated for each bin.

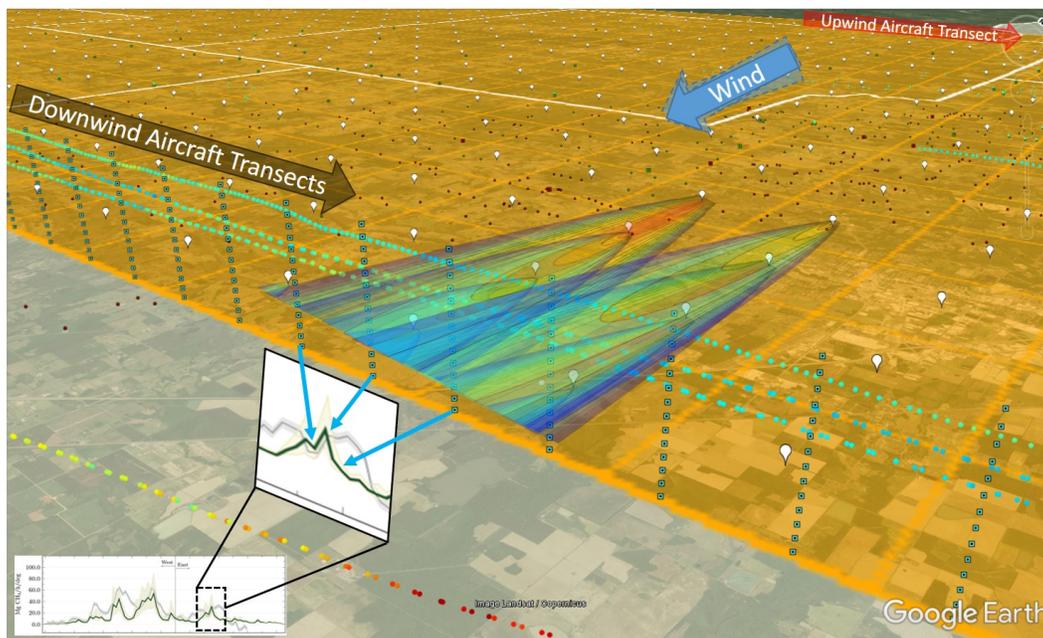


Figure 4.12: Emission estimates for each grid cell in the ground-level area estimate (GLAE) where idealized as point sources (white dots) and propagated downwind using Gaussian dispersion. Emission contributions from each grid cell are combined with daily average wind speed and direction to create simulated downwind transects (inset) comparable to longitudinal emission profiles measured by the aircraft. Actual aircraft measurement points from Schwietzke et al.²³ shown as small dots.

The structure of the simulated downwind transects show spatial features similar to the transects measured by the aircraft, as shown in Figure 4.13, providing confidence that both the aircraft and the bottom-up GLAEs are reasonably capturing emissions from the study area at a given point in time. Simulated downwind profiles were also calculated on an hourly basis, and qualitatively “best match” during the mid-day hours of the aircraft mass balance flights (see Appendix C.3).

However, despite strong agreement, the means of GLAEs calculated during flight times still differ significantly from aircraft results, as shown in Figure 4.11. Several possible explanations for disagreement have been hypothesized in the prior literature, most notably that super-emitters were inadequately characterized for production and other facility types^{69,96}. The GLAE model contains no external assumptions about the distribution or frequency of

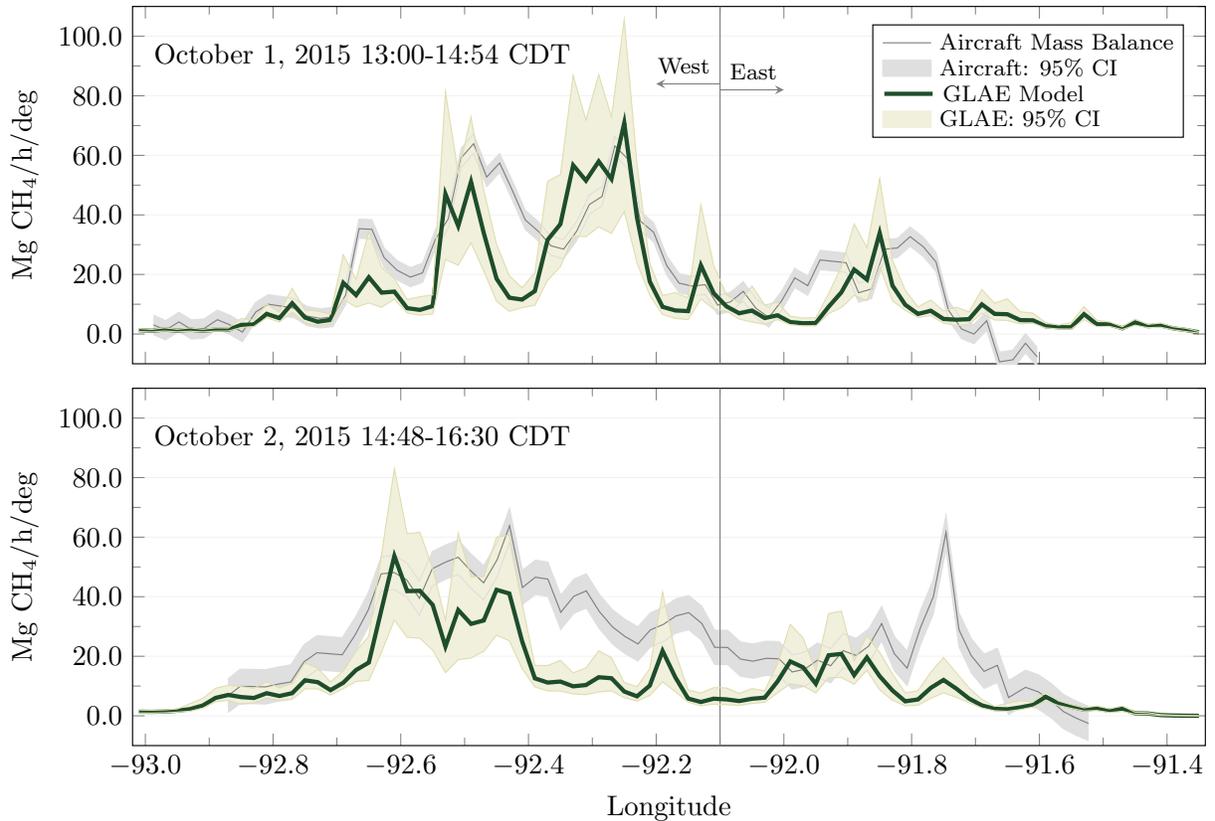


Figure 4.13: Simulated downwind transects produce spatial features similar to transects measured by the aircraft²³, providing confidence that both estimates are reasonably capturing study area emissions.

super-emitters; the probability of encountering a super-emitter is simulated at the frequency derived from random sampling during the field campaign. We therefore contend that modeling of super-emitters is not likely the cause of the remaining disagreement. Instead, we posit that a combination of sub-hourly timing of large magnitude events and a low-bias in the modeled MLU emission rate distribution could easily explain the difference in the means of top-down and bottom-up results. To better understand possible causes of disagreement in mean estimates, we consider four possible causes for the difference and perform sensitivity studies to assess their likelihood: (1) the modeled MLU emission rates, (2) gathering sector emissions, (3) non-oil and gas emissions, and (4) possible impacts from short duration event captured during aircraft transects.

Manual Liquid Unloading Emission Rate

First, since MLUs are the dominant contributor to modeled mid-day emissions, we increase modeled MLUs emission rates by 37% to match the aircraft mass balance study area total to within 1%. Recall that MLUs are modeled by drawing from a distribution that varies by a factor of five (247–1253 kg/h, mean = 513 kg/h) and that one MLU measured by tracer in this study was well above the mean of the modeled distribution (810 (603–1018 kg/h 95%CI)). Thus, a 37% increase in the mean MLU emission rate (i.e. 703 kg/h vs 513 kg/h) is plausible. Results from this analysis show near perfect agreement between top-down and bottom-up estimates on October 1st for total study area emissions, and substantially improved agreement for both eastern and western portions, as shown in Figure 4.14, and Appendix C.2.1.

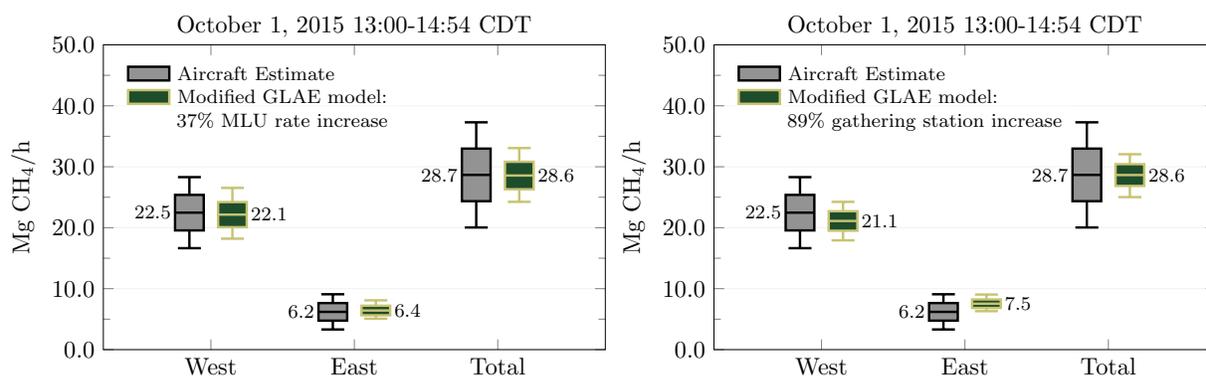


Figure 4.14: Increasing modeled manual liquid unloading (MLU) emission rates by 37% results in exceptional agreement between top-down and bottom-up estimates for the total study area, and both the eastern and western portions (left). Increasing modeled emission rates from gathering stations by 89% also results in total study area agreement, but poorer eastern and western apportionment (right).

Baseline Emission Estimate Errors from Gathering Stations

Emissions from gathering stations were also adjusted to match the aircraft mass balance study area total to within 1%, which required an 89% increase from the base scenario. In the production MLU adjustment scenario, the bottom-up estimate for the eastern portion of the study area was 4% greater than the aircraft estimate, while the western portion was 2% less than the aircraft estimate. In the gathering station adjustment scenario, the bottom-up estimate for the eastern portion of the study area was 22% greater than the aircraft estimate, while the western portion was 6% less than the aircraft estimate, as shown in Figure 4.14, and Appendix C.2.2. An 89% increase in gathering station emissions would require greater than a 150% increase in combustion slip emissions, which is unlikely because combustion slip was simulated from 111 recent tests that used reference methods. A ten-fold increase in the frequency of “stuck dump valves” using modeled rates from the base scenario would also account for the difference; however, this increased frequency is not supported by field observations. We therefore conclude that gathering station emissions likely could not explain differences in mean estimates.

Baseline Emission Estimates from Non-oil-and-gas Sources

Additional adjusted emission scenarios were also tested by increasing study area livestock, wetland, and geologic seep emissions to match aircraft results to within 1%. Livestock required a four-fold increase, wetlands required greater than an eight-fold increase, and geologic seeps required greater than a six-fold increase over the base scenario. While not impossible, increases of these magnitudes seem unlikely. Results from each of these scenarios is shown in Appendix C.2.

Timing of Manual Liquid Unloadings and Other Short Duration Events

Observing the downwind longitudinal profile on October 2nd in Figure 4.13, most of the disagreement between top-down and bottom-up in the eastern portion of the study area is due to emissions in the prominent top-down feature near -91.75° longitude, which is not present in the bottom-up model. This feature alone contributes ~ 4 Mg/h to the top-down estimate and the difference between top-down and bottom-up emissions attributable to this feature is ~ 3 Mg/h. This feature could be explained by one or more short-duration high-emission rate sources such as a blow-down, liquid unloading, or compressor engine start. For example, ~ 55 kg of gas is released in ~ 2 minutes by gas powered pneumatic motors and pumps during a compressor engine start, producing an instantaneous emission rate of $\sim 1,650$ kg/h. Emission rates from gas pneumatic starter motors are consistent and well defined by upstream gas pressures, unlike MLUs which have unknown and variable instantaneous emission rates. Study partners provided SCADA records of compressor engine start times and locations, and the GLAE model was re-run to exactly capture a compressor engine start, producing the emission profile shown in Figure 4.15. The sub-hourly down-wind transect simulation shows a prominent feature due to the compressor engine start that is absent in the hourly simulation, and in the minutes preceding and following the start-up as shown in Appendix C.2.6. The aircraft would transect the feature in ~ 2.5 minutes, roughly the duration of the event. A perfectly timed interception of this compressor engine start plume by the aircraft would increase the total study area emissions calculated by the aircraft by $\sim 1,650$ kg/h. Bottom-up simulations, even at the hourly scale, do not capture the instantaneous emission rates from these types of sources. Sub-hourly simulations indicate that these types of sources can create longitudinal emission rate profiles similar to those observed by the aircraft. Therefore, while timing of short-duration events may not explain the total difference in means, aircraft measurement of one or more short-duration events during downwind transects could increase aircraft estimates significantly above hourly average emission rates considered in the baseline model.

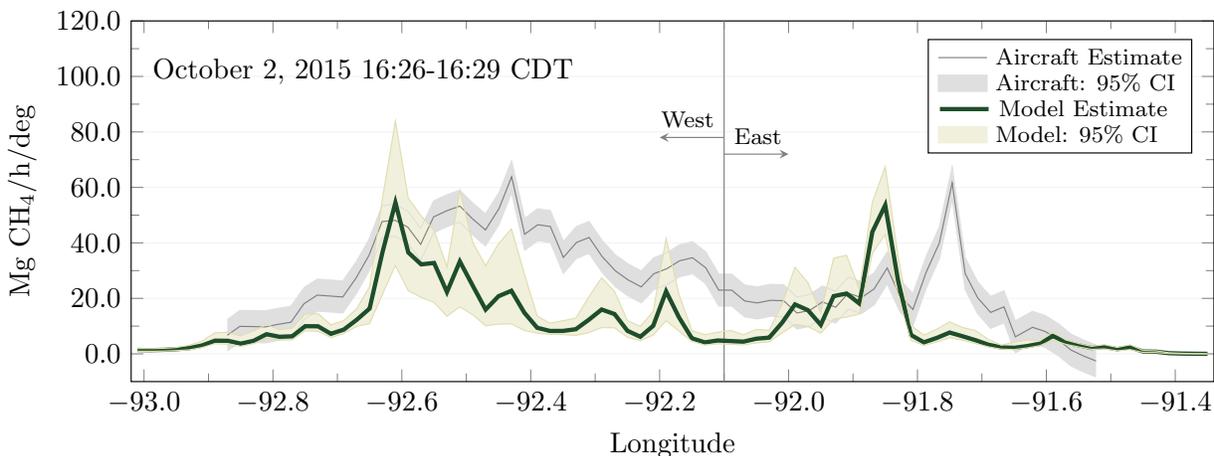


Figure 4.15: Short-duration, high-emission rate events such as blow-downs, liquid unloadings, and compressor engine gas starter motors can produce greatly enhanced local methane concentrations. The prominent feature in the model estimate between -91.8° and -91.9° longitude is due to a compressor engine start-up. This feature is absent in the minutes preceding and following the start-up as shown in Appendix C.2.6.

4.6 CONCLUSIONS

A detailed, spatially and temporally resolved, bottom-up inventory model was developed using Monte Carlo methods and compared to top-down aircraft mass balance flights on two consecutive days. Top-down and bottom-up estimates were found to agree (overlapping 95% confidence intervals) on both days, based on random sampling alone, by simply accounting for activity data at time-scales closer to the duration of aircraft mass balance flights. Yet, even with extensive contemporaneous emission measurements and highly specific activity data, the means of the estimates disagree. Some portion of disagreement may be due to other operators in the basin, unreported emissions, or inaccurate emission rate estimates in the bottom-up model. Plausible emission rate adjustments in one large source category alone could explain the difference completely, while maintaining consistent east and west apportionment between top-down and bottom-up estimates.

Comparing results from top-down and bottom methods directly without conditioning either or both models for time varying emissions is likely to result in disagreement. Aircraft mass balance flights deduce emission rates from instantaneous local atmospheric mixing

ratio enhancements and wind speed data measured on mid-day flights, during the work week. This timespan coincides with the peak of emissions from high-emission rate, short-duration sources initiated by workers. Thus, top-down aircraft mass balances produce larger emission rate estimates than bottom-up methods that average aggregated emissions over longer times. While both methods produce valid estimates for the timespans considered, they are not directly comparable without the temporal modeling presented here.

For aircraft measurements to provide valid annual estimates, a method that averages emissions over longer times, and accounts for short duration events is required. For bottom-up estimates to compare directly aircraft mass balance flights, detailed and specific activity data and emission factors are required for baseline and episodic sources. Baseline sources with steady emission rates, and intermittent sources that are small, numerous, and well-dispersed can be accounted for by typical bottom-up aggregation methods. The location, emission rate, and timing of blow-downs, liquid unloadings, compressor engine starts, and other short-duration, high-emission rate events capable of creating significant local atmospheric mixing ratio enhancements must be known exactly.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

Three studies were conducted to quantify CH₄ emissions from natural gas infrastructure. The first study, presented in Chapter 2, developed the first national CH₄ emission estimates from gathering facilities and processing plants. The second study, presented in Chapter 3, provided the first comparisons of facility-level CH₄ emissions from gathering stations utilizing both direct and atmospheric (downwind) measurement methods. Concurrent on-site and downwind tracer measurements were compared, and contemporaneous on-site and aircraft spiral flight measurements were compared. The third study, presented in Chapter 4, provided the first spatially and temporally resolved, top-down to bottom-up regional CH₄ emissions comparison to simulate emissions at hourly time scales based on hourly activity data corresponding to the flight times of aircraft mass balance flights.

In addition to the conclusions presented in each of the preceding chapters, these studies revealed several recommended practices, and potential areas of improvement for future measurements and estimates of CH₄ emissions from natural gas infrastructure.

- The importance of participation and involvement by industry “study partners” cannot be understated. The interpretation of measurement results would be much more difficult without site access, and the accompanying insight into operations provided by being on-site with the people who operate the measured facilities on a day to day basis. In certain regions, measurement campaigns that aimed to sample facilities without site access, using downwind methods from public roads, would have very few facilities available for measurement.
- Stack test data provided by study partners indicate that combustion slip emissions may vary significantly between engines within commonly used classifications for EPA emission factors. For example, extensive measurements from 2 models of four-stroke lean-burn natural gas engines showed emission rates that varied by a factor of 2. This

should be considered when comparing facility scale measurements where combustion slip contribute a significant portion of overall emissions.

- Releasing tracer gases from within exhaust stacks would help quantify the recovery rates of buoyant elevated plumes by tracer methods.
- Equipping the aircraft with tracer measurement capabilities would help distinguish emissions from the target facility from those of nearby sources.
- CH₄ emissions from dehydrator still vents may be underestimated by currently used software approved for estimating these emissions, based on a limited number of measurements obtained in this study. These emissions may be on the order of several kg/h, rather than several g/h as predicted by modeling software.
- Tank venting emissions at gathering station were observed infrequently, but were found to emit at very high rates that overshadowed emissions from all other sources at the facility. Identifying the frequency of occurrence, and duration of these emission would provide a better idea of there contribution to overall emission from gathering stations. Identifying the root cause of these emissions may allow for their avoidance.
- Manual liquid unloadings were shown to be a significant emission source in the production sector. An engineering solution to remove accumulated liquids from well bores, without the need to vent them to atmosphere could greatly reduce emissions from this source within the study area, and other producing regions where this activity is common practice.
- Detailed activity data, particularly for high-emission rate, episodic sources is critical for interpreting the results of measurements made at similar timescales. Top-down and bottom-up estimates were found to agree (overlapping 95% confidence intervals)

on consecutive days, based on unbiased random sampling alone, by simply accounting for activity data at time-scales closer to the duration of aircraft mass balance flights.

- An “intermediate scale” measurement technique would be extremely useful for quantifying emissions from major equipment where both direct measurement and downwind measurements are not possible. For example, tank venting emissions too large to measure using high-flow samplers, flow meters, or calibrated bags, and without suitable downwind road access for tracer vehicles.

REFERENCES

1. EIA. *U.S. Natural Gas Gross Withdrawals and Production* <https://www.eia.gov/dnav/ng/ng_prod_sum_dcu_NUS_a.htm> (2017).
2. EIA. *EIA - Annual Energy Outlook 2017* <<https://www.eia.gov/outlooks/aeo/>> (2017).
3. EIA. *Natural Gas Generators Make up the Largest Share of Overall U.S. Generation Capacity - Today in Energy - U.S. Energy Information Administration (EIA)* <<https://www.eia.gov/todayinenergy/detail.php?id=30872>> (2017).
4. 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. *Proceedings of the National Academy of Sciences* **109**, 6435–6440. ISSN: 0027-8424, 1091-6490 (Apr. 24, 2012).
5. Myhre, G. *et al.* in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. *et al.*) 659–740 (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013). ISBN: ISBN 978-1-107-66182-0. <www.climatechange2013.org>.
6. Turner, A. J. *et al.* A Large Increase in U.S. Methane Emissions over the Past Decade Inferred from Satellite Data and Surface Observations. *Geophysical Research Letters* **43**, 2016GL067987. ISSN: 1944-8007 (Mar. 16, 2016).
7. Kirschke, S. *et al.* Three Decades of Global Methane Sources and Sinks. *Nature Geoscience* **6**, 813–823. ISSN: 1752-0894 (Oct. 2013).
8. US EPA. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2014* (Washington, D.C., Apr. 15, 2016). <<https://www.epa.gov/ghgemissions/us-greenhouse-gas-inventory-report-1990-2014>>.
9. US EPA. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2014, Annex 3* EPA 430-R-16-002 (Washington, D.C., Apr. 15, 2016). <<https://www3.epa.gov/climatechange/ghgemissions/usinventoryreport.html>>.
10. AGA. *American Gas Association* <<https://www.aga.org/>> (2017).
11. EIA. *US Natural Gas Flow 2016* <<https://www.eia.gov/totalenergy/data/monthly/index.php>> (2017).
12. Matthew R. Harrison, Theresa M. Shires, Jane K. Wessels & R. Michael Cowgill. *Methane Emissions from the Natural Gas Industry* GRI-94/0257 and EPA-600/R96-080 (Gas Research Institute and U.S. Environmental Protection Agency, Washington, D.C., 1996).
13. Allen, D. T. *et al.* Measurements of Methane Emissions at Natural Gas Production Sites in the United States. *Proceedings of the National Academy of Sciences* **110**, 17768–17773. ISSN: 0027-8424, 1091-6490 (Oct. 29, 2013).

14. Mitchell, A. L. *et al.* Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results. *Environmental Science & Technology* **49**, 3219–3227. ISSN: 0013-936X (Mar. 3, 2015).
15. Marchese, A. J. *et al.* Methane Emissions from United States Natural Gas Gathering and Processing. *Environmental Science & Technology* **49**, 10718–10727. ISSN: 0013-936X (Sept. 1, 2015).
16. Roscioli, J. R. *et al.* Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Methods. *Atmos. Meas. Tech.* **8**, 2017–2035. ISSN: 1867-8548 (May 7, 2015).
17. 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. *Environmental Science & Technology* **49**, 3252–3261. ISSN: 0013-936X (Mar. 3, 2015).
18. Zimmerle, D. J. *et al.* Methane Emissions from the Natural Gas Transmission and Storage System in the United States. *Environmental Science & Technology*. ISSN: 0013-936X. doi:10.1021/acs.est.5b01669. <<http://dx.doi.org/10.1021/acs.est.5b01669>> (2015) (July 21, 2015).
19. Lamb, B. K. *et al.* Direct Measurements Show Decreasing Methane Emissions from Natural Gas Local Distribution Systems in the United States. *Environmental Science & Technology* **49**, 5161–5169. ISSN: 0013-936X (Apr. 21, 2015).
20. Lamb, B. K. *et al.* Development of Atmospheric Tracer Methods To Measure Methane Emissions from Natural Gas Facilities and Urban Areas. *Environmental Science & Technology* **29**, 1468–1479. ISSN: 0013-936X (June 1, 1995).
21. Robertson, A. M. *et al.* Variation in Methane Emission Rates from Well Pads in Four Oil and Gas Basins with Contrasting Production Volumes and Compositions. *Environmental Science & Technology*. ISSN: 0013-936X. doi:10.1021/acs.est.7b00571. <<http://dx.doi.org/10.1021/acs.est.7b00571>> (2017) (June 19, 2017).
22. Conley, S. *et al.* Application of Gauss’s Theorem to Quantify Localized Surface Emissions from Airborne Measurements of Wind and Trace Gases. *Atmos. Meas. Tech. Discuss.* **2017**, 1–29. ISSN: 1867-8610 (Apr. 18, 2017).
23. Schwietzke, S. *et al.* Improved Mechanistic Understanding of Natural Gas Methane Emissions from Spatially Resolved Aircraft Measurements. *Environmental Science & Technology*. ISSN: 0013-936X. doi:10.1021/acs.est.7b01810. <<http://dx.doi.org/10.1021/acs.est.7b01810>> (2017) (May 26, 2017).
24. 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. *Journal of Geophysical Research: Atmospheres* **119**, 2014JD021551. ISSN: 2169-8996 (June 27, 2014).
25. Ross, S. M. *Simulation* 326 pp. ISBN: 978-0-12-415825-2 (Academic Press, 2013).
26. EIAProcessing. *U.S. Natural Gas Plant Processing* <https://www.eia.gov/dnav/ng/ng_prod_pp_dcu_nus_a.htm> (2015).

27. OGJ. *SurveyDownloads* <<http://www.ogj.com/ogj-survey-downloads.html>> (2015).
28. Lyon, D. R. *et al.* Constructing a Spatially Resolved Methane Emission Inventory for the Barnett Shale Region. *Environmental Science & Technology* **49**, 8147–8157. ISSN: 0013-936X (July 7, 2015).
29. Theresa M. Shires & Matthew R. Harrison. *Methane Emissions from the Natural Gas Industry, Volume 6: Vented and Combustion Source Summary* GRI-94/0257.23 and EPA-600/R-96-080f (Gas Research Institute and U.S. Environmental Protection Agency, Washington, D.C., June 1996).
30. *Natural Gas Plant Processing* (United States Energy Information Agency, Sept. 30, 2014). <http://www.eia.gov/dnav/ng/ng_prod_pp_dcu_nus_a.htm> (2015).
31. EIA. *Natural Gas Annual 2015 (NGA) - Energy Information Administration - With Data for 2015* <<https://www.eia.gov/naturalgas/annual/>> (2015).
32. US EPA. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2012* EPA 430-R-14-003 (Washington, D.C., Apr. 15, 2014). <<http://www.epa.gov/climatechange/ghgemissions/usinventoryreport/archive.html>>.
33. Zimmerle, D. *et al.* (Submitted) Fayetteville Shale Methane Emissions from Gathering Pipelines and Scoping Guidelines for Future Pipeline Measurement Campaigns. *Elem Sci Anth* (2017).
34. Brandt, A. R., Heath, G. A. & Cooley, D. Methane Leaks from Natural Gas Systems Follow Extreme Distributions. *Environmental Science & Technology*. ISSN: 0013-936X. doi:10.1021/acs.est.6b04303. <<http://dx.doi.org/10.1021/acs.est.6b04303>> (2016) (Oct. 14, 2016).
35. Littlefield, J. A., Marriott, J., Schivley, G. A. & Skone, T. J. Synthesis of Recent Ground-Level Methane Emission Measurements from the U.S. Natural Gas Supply Chain. *Journal of Cleaner Production*. ISSN: 0959-6526. doi:10.1016/j.jclepro.2017.01.101. <<https://www.sciencedirect.com/science/article/pii/S0959652617301166>> (2017) (2017).
36. 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. *Environmental Science & Technology* **48**, 14508–14515. ISSN: 0013-936X (Dec. 16, 2014).
37. Yacovitch, T. I. *et al.* (Submitted) Natural Gas Facility Emission Measurements by Dual Tracer Flux in Two US Natural Gas Producing Basins. *Elem Sci Anth* (2017).
38. Conley, S. *et al.* (Under Review) Application of Gauss's Theorem to Quantify Localized Surface Emissions from Airborne Measurements of Wind and Trace Gases. *Atmos. Meas. Tech. Discuss.* **2017**, 1–29. ISSN: 1867-8610 (Apr. 18, 2017).
39. Karion, A. *et al.* Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region. *Environmental Science & Technology* **49**, 8124–8131. ISSN: 0013-936X (July 7, 2015).

40. Pétron, G. *et al.* Hydrocarbon Emissions Characterization in the Colorado Front Range: A Pilot Study. *Journal of Geophysical Research: Atmospheres* **117**, D04304. ISSN: 2156-2202 (D4 Feb. 27, 2012).
41. Howard, T., Ferrara, T. W. & Townsend-Small, A. Sensor Transition Failure in the High Flow Sampler: Implications for Methane Emission Inventories of Natural Gas Infrastructure. *Journal of the Air & Waste Management Association* **65**, 856–862. ISSN: 1096-2247 (July 3, 2015).
42. Allen, D. T., Sullivan, D. W. & Harrison, M. Response to Comment on “Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Pneumatic Controllers”. *Environmental Science & Technology* **49**, 3983–3984. ISSN: 0013-936X (Mar. 17, 2015).
43. Bacharach, Inc. *HI FLOW® Sampler For Natural Gas Leak Rate Measurement* July 2015. <www.mybacharach.com/wp-content/uploads/2015/08/0055-9017-Rev-7.pdf> (2016).
44. *GRI-GLYCalc Version 4.0* <<http://sales.gastechnology.org/000102.html>> (2017).
45. US EPA. *Method 19 - Sulfur Dioxide Removal and Particulate, Sulfur Dioxide and Nitrogen Oxides from Electric Utility Steam Generators* <<https://www.epa.gov/emc/method-19-sulfur-dioxide-removal-and-particulate-sulfur-dioxide-and-nitrogen-oxides-electric>> (2017).
46. US EPA. *Method 320 - Vapor Phase Organic and Inorganic Emissions by Extractive FTIR* <<https://www.epa.gov/emc/method-320-vapor-phase-organic-and-inorganic-emissions-extractive-ftir>> (2017).
47. Johnson, D. R., Covington, A. N. & Clark, N. N. Methane Emissions from Leak and Loss Audits of Natural Gas Compressor Stations and Storage Facilities. *Environmental Science & Technology* **49**, 8132–8138. ISSN: 0013-936X (July 7, 2015).
48. US EPA. *AP 42 Section 3.2 Natural Gas-Fired Reciprocating Engines - Related Information—Technology Transfer Network — Emissions Factors and Policy Applications Center— US EPA* <<http://www.epa.gov/ttn/chief/ap42/ch03/related/c03s02.html>>.
49. 40 C.F.R. § 98.33. *MANDATORY GREENHOUSE GAS REPORTING Subpart C: General Stationary Fuel Combustion Sources* <<http://www.ecfr.gov/>> (2017).
50. Myers, D. *Methane Emissions from the Natural Gas Industry, Volume 14: Glycol Dehydrators Final Report* EPA 600/R-96-080n (June 1996).
51. Caterpillar. *Caterpillar Application & Installation Guide Crankcase Ventilation Systems* <s7d2.scene7.com/is/content/Caterpillar/CM20160713-53120-62603> (2017).
52. US EPA. *Optimize Glycol Circulation and Install Flash Tank Separators in Glycol Dehydrators* <<https://www.epa.gov/natural-gas-star-program/optimize-glycol-circulation-and-install-flash-tank-separators-glycol>> (2017).

53. Martin Bland, J. & Altman, D. Statistical Methods for Assessing Agreement Between Two Methods Of Clinical Measurement. *The Lancet. Originally published as Volume 1, Issue 8476* **327**, 307–310. ISSN: 0140-6736 (Feb. 8, 1986).
54. Neri, F., Saitta, G. & Chiofalo, S. An Accurate and Straightforward Approach to Line Regression Analysis of Error-Affected Experimental Data. *Journal of Physics E: Scientific Instruments* **22**, 215 (1989).
55. Hollis, S. Analysis of Method Comparison Studies. *Annals of Clinical Biochemistry* **33** (Pt 1), 1–4. ISSN: 0004-5632 (Jan. 1996).
56. Draper, N. R. & Smith, H. *Applied Regression Analysis* 746 pp. ISBN: 978-0-471-17082-2 (Wiley, Apr. 23, 1998).
57. York, D. Least-Squares Fitting of a Straight Line. *Canadian Journal of Physics* **44**, 1079–1086. ISSN: 0008-4204 (May 1, 1966).
58. Altman, D. & Bland, J. Measurement in Medicine: The Analysis of Method Comparison Studies. *Journal of the Royal Statistical Society. Series D (The Statistician)* **32**, 307–317 (Sept. 1983).
59. Brandt, A. R. *et al.* Methane Leaks from North American Natural Gas Systems. *Science* **343**, 733–735. ISSN: 0036-8075, 1095-9203 (Feb. 14, 2014).
60. Harriss, R. *et al.* Using Multi-Scale Measurements to Improve Methane Emission Estimates from Oil and Gas Operations in the Barnett Shale Region, Texas. *Environmental Science & Technology* **49**, 7524–7526. ISSN: 0013-936X (July 7, 2015).
61. Lamb, B. K. *et al.* Direct and Indirect Measurements and Modeling of Methane Emissions in Indianapolis, Indiana. *Environmental Science & Technology* **50**, 8910–8917. ISSN: 0013-936X (Aug. 16, 2016).
62. Karion, A. *et al.* Methane Emissions Estimate from Airborne Measurements over a Western United States Natural Gas Field. *Geophysical Research Letters* **40**, 4393–4397. ISSN: 1944-8007 (Aug. 28, 2013).
63. Pétron, G. *et al.* 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**, 2013JD021272. ISSN: 2169-8996 (June 16, 2014).
64. Peischl, J. *et al.* Quantifying Atmospheric Methane Emissions from Oil and Natural Gas Production in the Bakken Shale Region of North Dakota. *Journal of Geophysical Research: Atmospheres* **121**, 2015JD024631. ISSN: 2169-8996 (May 27, 2016).
65. Miller, S. M. *et al.* Anthropogenic Emissions of Methane in the United States. *Proceedings of the National Academy of Sciences* **110**, 20018–20022. ISSN: 0027-8424, 1091-6490 (Oct. 12, 2013).
66. Peischl, J. *et al.* Quantifying Atmospheric Methane Emissions from the Haynesville, Fayetteville, and Northeastern Marcellus Shale Gas Production Regions. *Journal of Geophysical Research: Atmospheres* **120**, 2014JD022697. ISSN: 2169-8996 (Mar. 16, 2015).

67. Heath, G. A., O'Donoughue, P., Arent, D. J. & Bazilian, M. Harmonization of Initial Estimates of Shale Gas Life Cycle Greenhouse Gas Emissions for Electric Power Generation. *Proceedings of the National Academy of Sciences* **111**, E3167–E3176. ISSN: 0027-8424, 1091-6490 (May 8, 2014).
68. Zavala-Araiza, D. *et al.* Toward a Functional Definition of Methane Super-Emitters: Application to Natural Gas Production Sites. *Environmental Science & Technology* **49**, 8167–8174. ISSN: 0013-936X (July 7, 2015).
69. Zavala-Araiza, D. *et al.* Super-Emitters in Natural Gas Infrastructure Are Caused by Abnormal Process Conditions. *Nature Communications* **8**, 14012. ISSN: 2041-1723 (Jan. 16, 2017).
70. Allen, D. T. Methane Emissions from Natural Gas Production and Use: Reconciling Bottom-up and Top-down Measurements. *Current Opinion in Chemical Engineering. Energy and environmental engineering / Reaction engineering* **5**, 78–83. ISSN: 2211-3398 (Aug. 2014).
71. Bell, C. S. *et al.* (Submitted) Comparison of Methane Emission Estimates from Multiple Measurement Techniques at Natural Gas Production Pads. *Elem Sci Anth* (2017).
72. Allen, D. T. *et al.* Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Liquid Unloadings. *Environmental Science & Technology* **49**, 641–648. ISSN: 0013-936X (Jan. 6, 2015).
73. Allen, D. T. *et al.* Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Pneumatic Controllers. *Environmental Science & Technology* **49**, 633–640. ISSN: 0013-936X (Jan. 6, 2015).
74. Vaughn, T. L. *et al.* (Submitted) Comparing Facility-Level Methane Emission Rate Estimates at Natural Gas Gathering and Boosting Stations. *Elem Sci Anth* (2017).
75. Of Air Quality Planning, O. & Standards, U. E. *Emissions Factors & AP 42* <<http://www.epa.gov/ttn/chief/ap42/index.html>>.
76. 40 C.F.R. § 98.230. *MANDATORY GREENHOUSE GAS REPORTING Subpart W: Petroleum and Natural Gas Systems* <<http://www.ecfr.gov/>> (2017).
77. Pickering, C. *METHANE EMISSIONS FROM GATHERING PIPELINE NETWORKS, DISTRIBUTION SYSTEMS, AGRICULTURE, WASTE MANAGEMENT AND NATURAL SOURCES* Masters Thesis (Colorado State University, Fort Collins, CO, 2016).
78. GTI. *Field Measurement Program to Improve Uncertainties for Key Greenhouse Gas Emission Factors for Distribution Sources* OTD-10-0002 (Nov. 10, 2009).
79. IPCC 2006. *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme* (eds Eggleston H.S., Buendia L., Miwa K., Ngara T. & Tanabe K.) 2006. <<http://www.ipcc-nggip.iges.or.jp/public/2006gl/vol4.html>>.
80. *USDA - NASS, Census of Agriculture - 2012 Census Volume 1, Chapter 2: County Level Data* <https://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1,_Chapter_2_County_Level/Arkansas/> (2017).

81. Smartt, A. D. *et al.* Chamber Size Effects on Methane Emissions from Rice Production. *Open Journal of Soil Science* **05**, 227–235. ISSN: 2162-5360, 2162-5379 (2015).
82. U.S. Fish & Wildlife Service. *National Wetlands Inventory: State Downloads* <<https://www.fws.gov/wetlands/Data/State-Downloads.html>> (2017).
83. Bartlett, K. B., Crill, P. M., Bonassi, J. A., Richey, J. E. & Harriss, R. C. Methane Flux from the Amazon River Floodplain: Emissions during Rising Water. *Journal of Geophysical Research: Atmospheres* **95**, 16773–16788. ISSN: 2156-2202 (D10 Sept. 20, 1990).
84. Deemer, B. R. *et al.* Greenhouse Gas Emissions from Reservoir Water Surfaces: A New Global Synthesis. *BioScience* **66**, 949–964. ISSN: 0006-3568 (Nov. 1, 2016).
85. *Water Quality Monitoring Data — ADEQ* <https://www.adeq.state.ar.us/techsvs/env_multi_lab/water_quality_station.aspx> (2017).
86. Holgerson, M. A. & Raymond, P. A. Large Contribution to Inland Water CO₂ and CH₄ Emissions from Very Small Ponds. *Nature Geoscience* **9**, 222–226. ISSN: 1752-0894 (Mar. 2016).
87. Bastviken, D., Tranvik, L. J., Downing, J. A., Crill, P. M. & Enrich-Prast, A. Freshwater Methane Emissions Offset the Continental Carbon Sink. *Science* **331**, 50–50. ISSN: 0036-8075, 1095-9203 (Jan. 7, 2011).
88. Etiope, G. & Klusman, R. W. Microseepage in Drylands: Flux and Implications in the Global Atmospheric Source/Sink Budget of Methane. *Global and Planetary Change. Quaternary and Global Change: Review and Issues Special issue in memory of Hugues FAURE* **72**, 265–274. ISSN: 0921-8181 (July 2010).
89. *Population Estimates, July 1, 2016, (V2016)* <<http://www.census.gov/quickfacts/>> (2017).
90. *Integrated Database: Arkansas Sewer Data* <http://www.nesc.wvu.edu/septic_idb/arkansas.htm#septicstats> (2017).
91. Leverenz, H. L., Tchobanoglous, G. & Darby, J. *Evaluation of Greenhouse Gas Emissions from Septic Systems* in collab. with IWA Publishing & Water Environment Research Foundation. OCLC: ocn890654374. 1 p. ISBN: 978-1-84339-616-1 (Water Environment Research Foundation ; IWA Publishing, Alexandria, VA : London, UK, 2010).
92. *EPA Facility Level GHG Emissions Data* <<http://ghgdata.epa.gov/ghgp/main.do>> (2017).
93. AOGC. *Arkansas Oil & Gas Commission* <<http://www.aogc.state.ar.us/default.aspx?aspxerrorpath=/default.aspx>> (2017).
94. US EPA, O. *GHG Reporting Program Data Sets* <<https://www.epa.gov/%5C%0020ghgreporting/ghg-reporting-program-data-sets>> (2017).
95. Turner, D. B. *Workbook of Atmospheric Dispersion Estimates* (1970).
96. Zavala-Araiza, D. *et al.* Reconciling Divergent Estimates of Oil and Gas Methane Emissions. *Proceedings of the National Academy of Sciences* **112**, 15597–15602. ISSN: 0027-8424, 1091-6490 (Dec. 22, 2015).

97. *Numerical Recipes in C: The Art of Scientific Computing* (ed Press, W. H.) 2nd ed (Cambridge University Press, Cambridge ; New York, 1992). 994 pp.

APPENDIX A

ESTIMATION OF METHANE EMISSIONS FROM THE U.S. NATURAL GAS GATHERING AND PROCESSING SECTOR

Facility Number	Facility Type	Sampling Date	State	Facility Natural Gas Throughput						FLER ($\frac{\text{kg}}{\text{hr}}$ CH ₄)	Unbiased Weighted Std Dev ($\frac{\text{kg}}{\text{hr}}$ CH ₄)	tnFLER (% of CH ₄ throughput)
				(MMscfd)	CH ₄ (%)	C ₂ H ₆ (%)	C ₃ H ₈ (%)	CO ₂ (%)				
1	C	2/6/2014	TX	37	77.1	12.9	5.3	1	255.1	84.4	1.11%	
2	C	4/8/2014	UT	20	90.7	5.1	1.8	0.9	94.3	60	0.65%	
3	C	2/13/2014	TX	0.6	91.2	4.1	1	1.5	74.5	16.2	17.52%	
4	C	2/13/2014	TX	-	96.3	1.9	0.4	0.8	64.8	56.1	-	
5	C	4/9/2014	UT	20	87.1	5.8	1.9	1	62.4	17.7	0.45%	
6	C	4/9/2014	UT	36	90.7	5.1	1.8	0.9	43.9	23.7	0.17%	
7	C	11/6/2013	WY	6.4	82.7	10.8	3.3	0.8	43.8	8.4	1.03%	
8	C	4/7/2014	UT	6.3	90.7	5.1	1.8	0.9	41.6	30.9	0.90%	
9	C	1/29/2014	TX	10.5	76.3	14	5.5	1.2	29.6	9.7	0.46%	
10	C	2/6/2014	TX	14.7	70.3	14.8	7.3	2	28.8	6.5	0.35%	
11	C	12/4/2013	PA	607	97.7	1.9	0.1	-	21.9	6.3	< 0.01%	
12	C	3/25/2014	NM	59	86.7	6.4	2.7	1.8	19.8	3.1	0.05%	
13	C	1/28/2014	TX	6.3	73.5	14.3	6	1	17.9	3.2	0.48%	
14	C	2/12/2014	TX	0.2	96.9	0.3	0	2.2	14.8	2.1	8.45%	
15	C	2/19/2014	KS	2.8	73.8	6.8	4.3	0.1	13.5	2.6	0.81%	
16	C	4/14/2014	CO	20.2	77.3	11.5	4.5	2.5	9.3	8.2	0.07%	
17	C	3/17/2014	OK	6.2	85.3	7.7	3.5	0.3	9.2	3	0.22%	
18	C	2/18/2014	KS	3.3	80.7	5.4	3	0.2	8.4	2	0.39%	
19	C	3/25/2014	NM	2	86.1	6.6	2.9	1.5	8.3	6.2	0.60%	
20	C	4/1/2014	WY	2.2	81.8	0.4	-	16.9	8.3	2.4	0.58%	
21	C	3/24/2014	OK	0.7	90.1	4.4	1.9	0.1	7.9	1.9	1.58%	
22	C	2/14/2014	TX	0.6	95.8	0.5	0	3.3	7.7	9.4	1.66%	
23	C	2/18/2014	KS	2.8	77.2	5.1	2.8	0.2	7.1	0.9	0.41%	
24	C	11/6/2013	WY	6.8	81.1	11.8	4.3	0.6	6.2	1.5	0.14%	
25	C	2/10/2014	TX	0.3	97.5	0.5	0.1	1.4	5.8	1.4	2.25%	
26	C	2/19/2014	KS	3.8	73.4	6.7	4.3	0.2	5.5	1.3	0.24%	
27	C	2/19/2014	KS	4.5	75.4	6.4	3.7	0.1	5.2	1.3	0.19%	
28	C	2/13/2014	TX	-	95.6	2.6	0.5	0.5	5	6.3	-	
29	C	2/11/2014	TX	0.7	95.4	1.7	0.8	1	5	1.4	0.97%	
30	C	11/6/2013	WY	64.9	82.7	10.8	3.5	0.7	3.4	0.4	< 0.01%	
31	C	3/18/2014	OK	1	90.8	3.8	2	0.3	2.7	3.2	0.38%	
32	C	3/24/2014	OK	-	77.7	10.3	6	0.4	1.6	1	-	
33	C	2/19/2014	KS	1.2	74.7	6.5	3.8	0.1	1.2	0.5	0.17%	
34	C	2/18/2014	KS	1.7	81	5.1	2.5	0.2	0.9	0.1	0.08%	
35*	C/D	10/25/2013	TX	9.5	95.7	0.9	0	2.7	698.6	63.2	9.54%	
36	C/D	3/27/2014	NM	42.5	92.7	0.6	0.1	6.3	344.1	95.2	1.09%	
37	C/D	2/26/2014	AR	32	95.2	1	0	3.4	240.5	66.9	0.98%	
38	C/D	2/6/2014	AR	26	93.5	1.1	0	5.1	217.6	79.4	1.12%	
39	C/D	2/24/2014	AR	65	96.8	1.1	0	1.7	196.1	64	0.39%	
40	C/D	2/11/2014	TX	13.1	76	11.9	6.9	0.2	195.3	67.8	2.44%	
41	C/D	12/2/2013	NY	206	97.7	1.9	0.1	0	159.8	71.2	0.10%	
42	C/D	3/26/2014	OK	35	88.2	5.5	2.4	0.2	146.7	67	0.59%	
43	C/D	3/26/2014	OK	28	85.3	6.9	3.1	0.2	145.2	29.4	0.76%	
44	C/D	2/26/2014	AR	44	94.7	1.1	0	3.9	119	86	0.36%	
45	C/D	4/3/2014	NM	23.5	79.5	1.2	0.3	18.9	112	58.7	0.75%	
46	C/D	12/3/2013	PA	291.5	97.5	2.1	0.1	0	109.1	98	0.05%	
47	C/D	2/5/2014	AR	37.5	96.7	1.3	0	1.6	106.6	56.3	0.37%	
48	C/D	2/25/2014	AR	47	97.7	1.4	0	0.7	98.6	72.7	0.27%	
49	C/D	2/5/2014	AR	29.6	97	1.3	0	1.5	93.4	19.1	0.41%	
50	C/D	3/20/2014	TX	28	78.8	10.5	5.2	0.5	92.8	41	0.52%	
51	C/D	2/11/2014	TX	10	76.6	11.1	7	0.3	82.9	33.2	1.35%	
52	C/D	2/24/2014	AR	30	96.3	1.3	0	2	76.3	32.2	0.33%	
53	C/D	12/13/2013	PA	10	95.8	2.5	0.3	0.2	76	39.2	0.99%	
54	C/D	3/28/2014	NM	22.4	79.3	0.9	0.2	19.6	71.7	17.1	0.50%	
55	C/D	10/24/2013	TX	60.3	95.5	1.5	0.1	1.9	66.7	32.9	0.14%	
56	C/D	10/23/2013	TX	22	95.6	1.6	0.1	1.6	63.5	51.1	0.38%	
57	C/D	3/31/2014	NM	47.6	92.5	1.7	0.4	5	61.1	17.4	0.17%	
58	C/D	2/3/2014	TX	23.5	79.3	11.7	5.1	0.9	59.4	20.9	0.40%	
59*	C/D	10/25/2013	TX	9.5	95.7	0.9	0	2.7	49.7	13.9	0.68%	
60	C/D	4/2/2014	WY	4	92.9	0	-	4.6	48.6	11.8	1.63%	
61	C/D	2/5/2014	TX	9.3	75.9	13.3	5.5	1.4	47.7	21.7	0.84%	
62	C/D	2/3/2014	AR	29	95.1	1.3	0	3.4	47.6	23.7	0.21%	

Continued on next page

Continued from previous page

Facility Number	Facility Type	Sampling Date	State	Facility Natural Gas					FLER ($\frac{\text{kg}}{\text{hr}}$ CH ₄)	Unbiased Weighted Std Dev ($\frac{\text{kg}}{\text{hr}}$ CH ₄)	tnFLER (% of CH ₄ throughput)
				Throughput (MMscfd)	CH ₄ (%)	C ₂ H ₆ (%)	C ₃ H ₈ (%)	CO ₂ (%)			
63	C/D	3/25/2014	OK	0.4	85.3	7.9	3.4	0.9	46.8	12.1	17.08%
64	C/D	12/5/2013	PA	511	97.2	2.4	0.1	0	45.1	18.9	0.01%
65	C/D	4/16/2014	CO	38	70.2	14.7	7.2	2.6	36.4	29.9	0.17%
66	C/D	12/12/2013	PA	9	96.4	2.4	0.2	0.2	35.8	8.9	0.51%
67	C/D	3/27/2014	OK	2.2	90.5	4.5	1.9	0.3	34.5	5.6	2.16%
68	C/D	3/27/2014	NM	16.6	84.5	0.6	0.1	14.7	34.5	12.9	0.31%
69	C/D	2/25/2014	AR	18	97	1.4	0	1.4	33.9	10	0.24%
70	C/D	3/20/2014	OK	15.5	83.8	8.9	3.4	0.5	32.7	22.4	0.31%
71	C/D	4/17/2014	CO	8.8	67.7	14	9.1	2.5	26.5	8.8	0.55%
72	C/D	3/27/2014	OK	5.2	92.8	4	1.5	0.2	26.4	12.9	0.68%
73	C/D	2/27/2014	AR	11	96.1	1.1	0	2.5	26.2	10.8	0.31%
74	C/D	4/1/2014	WY	3	92.3	0.1	-	4.4	26	10.4	1.17%
75	C/D	10/29/2013	WY	10	88.1	6.1	2.7	0.6	25.9	6.9	0.37%
76	C/D	2/7/2014	AR	28.2	96.6	1.2	0	1.7	23.8	12.2	0.11%
77	C/D	2/12/2014	TX	3.6	97.5	0.2	0	1.7	22.3	13.1	0.79%
78	C/D	3/19/2014	OK	0.3	87.3	5.4	2.2	0.8	22	1	11.84%
79	C/D	3/18/2014	OK	5.8	95.5	2.4	0.6	0.7	21.7	5.3	0.49%
80	C/D	2/14/2014	TX	0.7	97.7	0.4	0	1.4	17.9	6	3.04%
81	C/D	11/7/2013	WY	37.7	88.1	6.3	2.6	0.6	17.7	3.6	0.07%
82	C/D	4/2/2014	WY	6	70	13.3	7.9	1.5	15	5	0.45%
83	C/D	10/22/2013	TX	50.6	94.9	2.6	0.2	1.5	13.7	2	0.04%
84	C/D	2/14/2014	TX	2.9	96.5	1.1	0.2	1.5	13.2	5.4	0.59%
85	C/D	3/28/2014	NM	6.5	76.7	1.7	0.6	20.6	12.8	5.7	0.32%
86	C/D	3/19/2014	OK	4.5	85.3	7.4	3.2	0.8	11.5	4.3	0.37%
87	C/D	1/27/2014	TX	10	64.7	18.4	10	1.1	11.4	3.7	0.22%
88	C/D	12/11/2013	PA	5.4	95.2	3.2	0.5	0.1	11	3.7	0.27%
89	C/D	3/24/2014	NM	11.8	83.9	0.6	0.1	15.4	10.9	4.3	0.14%
90	C/D	12/11/2013	PA	120	97	2.2	0.1	0.4	9	2.5	< 0.01%
91	C/D	10/22/2013	TX	15.4	95.9	1.9	0.1	1.5	8.9	9.1	0.07%
92	C/D	12/16/2013	WV	0.5	74	16.3	5.8	0.2	8.7	8.4	2.92%
93	C/D	11/5/2013	WY	135	83.8	7.6	3.3	3	8.1	4.7	< 0.01%
94	C/D	2/3/2014	TX	0	75.9	13.2	5.7	1.3	7.5	4.2	69.60%
95	C/D	2/26/2014	AR	34	95.8	1	0	2.7	7.5	2.3	0.03%
96	C/D	2/10/2014	TX	0.3	97.9	0.4	0	1.5	6.1	1.4	2.52%
97	C/D	12/9/2013	PA	0.2	92.5	3.5	0.5	0.1	4.6	0.5	4.17%
98	C/D	4/16/2014	CO	1	67.5	13.9	8.1	2.5	2.4	0.6	0.45%
99	C/D	12/17/2013	WV	0.7	77	15.1	4.9	0.1	2	0.7	0.47%
100	C/D	3/17/2014	OK	-	84.3	4.9	3.2	0.1	1.4	0.5	-
101	C/D	3/21/2014	OK	-	93.4	3.3	1.1	1	0.7	0.2	-
102	C/D/T	2/4/2014	TX	60.1	76.5	13.2	5.7	1.1	240.5	148.8	0.65%
103	C/D/T	1/27/2014	TX	400	74.2	14.4	6.3	1	173.7	37.1	0.07%
104	C/D/T	2/12/2014	TX	19	76.3	11.5	7	0.2	142.1	87.1	1.16%
105	C/D/T	4/7/2014	UT	84.8	89.8	5.3	1.8	0.9	40.4	27.7	0.07%
106	C/D/T	4/9/2014	UT	62.4	89.8	5.3	1.8	0.9	34.1	17.2	0.08%
107	C/D/T	1/28/2014	TX	11	74.2	14.4	6.3	1	28.3	15.1	0.43%
108	C/D/T	1/28/2014	TX	12	74.2	14.4	6.3	1	20.1	5.6	0.28%
109	C/D/T	1/31/2014	TX	1.4	72.7	14.9	6.8	1.8	6.5	2.2	0.75%
110	D	11/5/2013	WY	8.5	80.7	8.7	4.3	2.8	38	10.1	0.69%
111	D	12/6/2013	PA	4.1	97.3	2.3	0.1	0	10.6	4.6	0.33%
112	D	12/17/2013	WV	0.6	76.1	15.9	5.1	0	7.8	2.3	2.07%
113	D	12/10/2013	PA	5	97.9	1.4	0.1	0.4	3.5	2.7	0.09%
114	D	12/4/2013	PA	-	97.4	2.2	0.1	0	1.9	0.6	-
115	D/T	3/26/2014	NM	320	88.4	0.9	0.2	9.4	142.4	49.7	0.06%
116	P	4/11/2014	CO	627	89.3	5.1	1.3	3.1	606	290.7	0.13%
117	P	2/20/2014	LA	197	86.5	6.3	4	0.7	451.1	191.9	0.29%
118	P	11/8/2013	WY	972.4	87.5	7.4	2.4	0.8	279.4	107.7	0.04%
119	P	4/1/2014	CO	429	89.5	3.6	1.5	4.1	267.7	141.3	0.08%
120	P	4/2/2014	NM	136	87.1	6.7	2.5	1.6	207.1	118.9	0.20%
121	P	3/31/2014	WY	41.9	61.9	15	10.9	1.5	166.6	102.2	0.80%
122	P	4/15/2014	CO	205	75.5	13.3	5.7	2.5	156.8	147.7	0.13%
123	P	3/20/2014	CO	155	69	14.4	6.1	2.6	128.2	65.7	0.11%

Continued on next page

Continued from previous page

Facility Number	Facility Type	Sampling Date	State	Facility Natural Gas					FLER ($\frac{\text{kg}}{\text{hr}}$ CH ₄)	Unbiased Weighted		tnFLER (% of CH ₄ throughput)
				Throughput (MMscfd)	CH ₄ (%)	C ₂ H ₆ (%)	C ₃ H ₈ (%)	CO ₂ (%)		Std Dev ($\frac{\text{kg}}{\text{hr}}$ CH ₄)		
124	P	10/30/2013	WY	205	88.1	6.1	2.7	0.6	112.5	9.5	0.08%	
125	P	12/18/2013	WV	337	80.6	13	3.9	0.2	93.2	68.5	0.04%	
126	P	11/4/2013	WY	614	83.2	7.9	3.4	3	75.5	26.1	0.02%	
127	P	4/10/2014	CO	485	90.8	5.5	1.4	1	58.4	21.8	0.02%	
128	P	2/17/2014	TX	266	84.5	8.5	3.5	0.9	54.9	21.9	0.03%	
129	P	2/21/2014	AL	210.3	91.1	4.3	2	0.9	39.4	24.1	0.02%	
130	P	1/30/2014	TX	193	75.7	13.3	5.8	1.1	14.1	7.2	0.01%	
131	P	4/3/2014	WY	2	58.9	12.1	13.9	1	3.3	1.3	0.35%	

* Facilities 35 and 59 are a single facility that was sampled in both a high emitting state and lower emitting state.

Note that 1 MMscfd = 0.328 m³/s.

APPENDIX B

COMPARING FACILITY-LEVEL METHANE EMISSION RATE ESTIMATES AT NATURAL GAS GATHERING AND BOOSTING STATIONS

B.1 ALTERNATE METHOD COMPARISONS USING SOEs DEVELOPED FROM MEASURED DEHYDRATOR REGENERATOR VENTS

This section reports results from alternate method comparisons that calculate simulated dehydrator regenerator vents based on 4 dehydrator units measured in this study. Additionally, compressor engine crankcase vents are calculated based on recent measurements by Johnson et al.⁴⁷ All other SOE categories are calculated in the same way as method comparisons presented in the main text.

B.1.1 SOE AND OVERALL RESULTS SUMMARY

Simulated Combustion Slip was the largest source category and contributed 63% to the cumulative SOE for the 17 facilities included in method comparisons shown in Figure B.1. *ODMs* contributed 14%, *SDMs* contributed 5%, *Simulated Crankcase Vents* contributed 7%, and *Simulated Dehydrator Regenerator Vents* contributed 10% to the cumulative SOE. For each measurement method, 95% confidence intervals indicate that the method would produce a FLER within the interval 95% of the time. We consider methods with overlapping confidence intervals to agree. Tracer and SOE 95% confidence intervals overlap at 10 out of 14 facilities, while aircraft and SOE confidence intervals overlap at five out of six facilities.

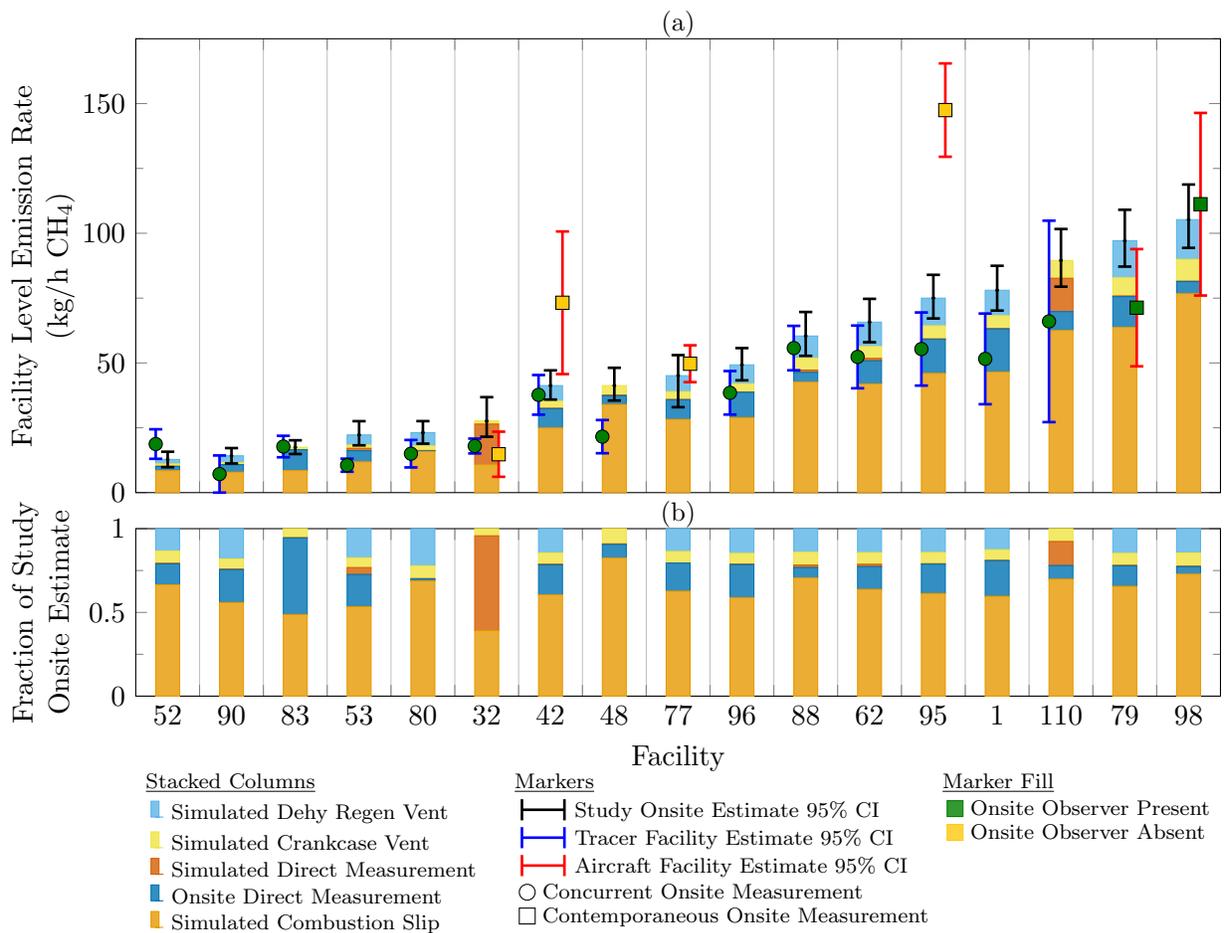


Figure B.1: Facility-level CH₄ emission rate summary at all facilities included in method comparisons. Study on-site estimates (SOE) are the sum of on-site direct measurements plus engineering estimates for unmeasured sources (stacked columns, black error bars). Tracer (left mark, blue error bars) and aircraft (right mark, red error bars) are overlaid at facilities where these measurements were compared to SOEs. Marker shape and fill indicate same/different day and the presence/absence of on-site observers, which influence the comparability of measurements. Bottom panel illustrates the fraction of the SOE contributed by each component; combustion slip contributes more than half of emissions at 15 of 17 facilities and accounts for two thirds of cumulative SOE emissions for these 17 facilities.

B.1.2 TRACER FACILITY ESTIMATE AND STUDY ON-SITE ESTIMATE COMPARISON

When compared in aggregate by difference plot and variance-weighted least-squares regressions, tracer predicts lower FLER than SOE for 14 concurrently-measured gathering stations at the 95% confidence level (see Figure B.2). In Figure B.2(a) the difference of tracer and SOE is plotted against the uncertainty weighted mean of tracer and SOE. The mean of differences (termed “bias”) is -10.9 kg/h, indicating that tracer predicts lower FLER than SOE. A paired t-test is used to determine if the bias is significant. The shaded area in Figure B.2(a) highlights the 95% confidence interval on bias. The confidence interval does not include $x = 0$, which indicates that the bias is statically significant at the 95% confidence level. The “limits of agreement” are given by two standard deviations of method differences and provide an assessment of method agreement based on the measured data. The limits of agreement for tracer and SOE are ± 17.6 kg/h (dash-dot lines in Figure B.2(a)), indicating that tracer may predict a FLER 28.5 kg/h less than or 6.7 kg/h greater than SOE.

In Figure B.2(b) a VWLS regression (dashed line) is performed on tracer and SOE. The slope of the regression (tracer = $0.76 \cdot \text{SOE}$, $R^2 = 0.92$) is less than unity, indicating that tracer predicts lower FLER than SOE. The 95% confidence interval (shaded region) on the regression slope (tracer = $0.69 \cdot \text{SOE}$ to tracer = $0.83 \cdot \text{SOE}$) does not include the line of equality ($y = x$), indicating that tracer predicts lower FLER than SOE at the 95% confidence level.

Table B.1 shows the contribution of SOE component categories to the cumulative SOE for facilities included in the tracer to SOE method comparison.

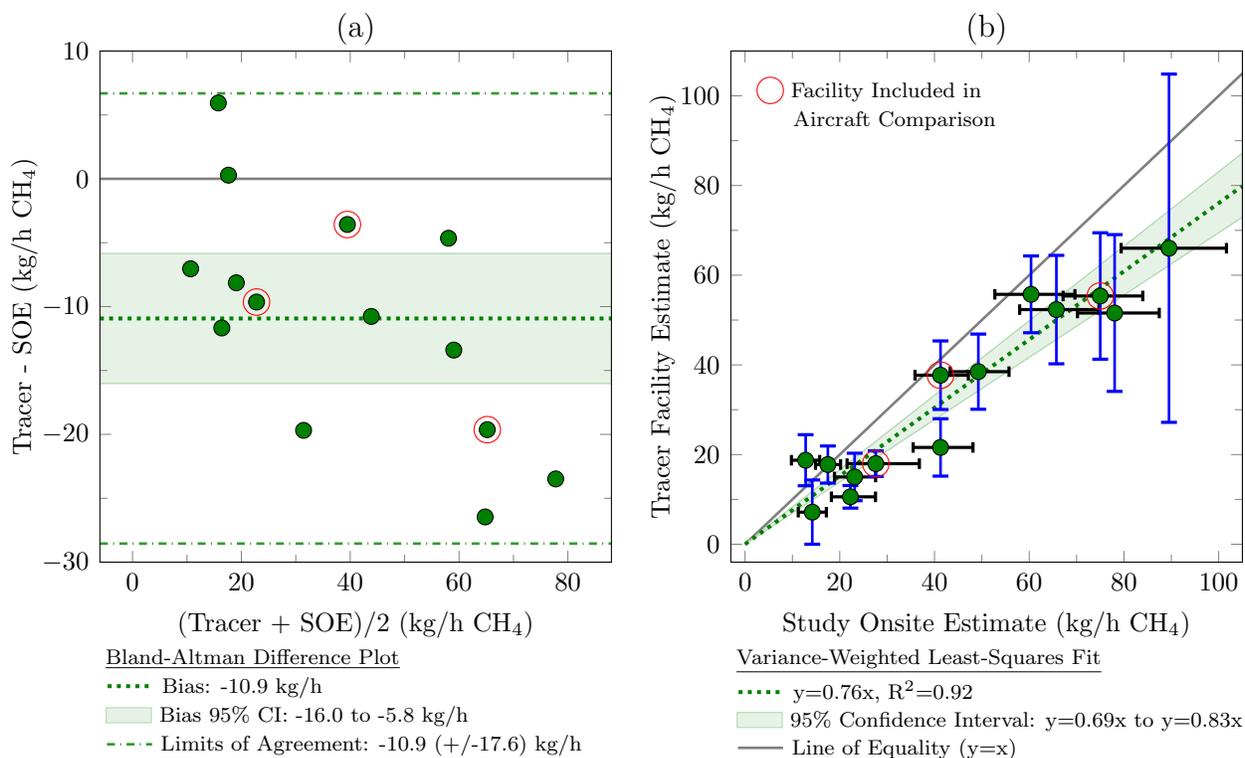


Figure B.2: Tracer predicts lower facility-level CH₄ emission rates than study on-site estimates at the 95% confidence level using difference plot.

Table B.1: Comparison of cumulative CH₄ emission rates for 14 gatherings stations included in the TFE/SOE comparison, showing all categories contributing to the SOE. Combustion slip is the largest contributor to the SOE.

CH ₄ Emission Source	Onsite Direct Measurements	Simulated Direct Measurements			Simulated
		Above Hi-Flow Range	Below Hi-Flow Range	Observed Not Measured	
Compressor Units	42.6	-	0.5	5.7	-
Pressure Relief Valves	3.4	12.3	-	-	-
Rod Packing Vents	27.3	12.3	0.2	-	-
Dehydrator Units	2.2	-	0.1	-	-
Other	4.5	-	0.3	0.4	-
Pig Launchers/Receivers	0.1	-	-	-	-
Piping or Gas Lines	0.4	-	0.2	-	-
Separators	1.7	-	-	-	-
Tanks	4.8	-	0.0	-	-
Combustion Slip	-	-	-	-	391.3
Crankcase Vents	-	-	-	-	42.9
Dehydrator Regenerator Vents	-	-	-	-	64.8
Cumulative Study Onsite Estimate	86.8	24.7	1.3	6.2	617.9 kg/h
Cumulative Tracer Facility Estimate					466.0 kg/h

B.1.3 AIRCRAFT FACILITY ESTIMATE AND STUDY ON-SITE ESTIMATE COMPARISON

Aircraft predicts higher FLER than SOE when compared by difference plot and VWLS regression, as shown in Figure B.3. When compared by difference plot, aircraft is biased high relative to SOE (16.9 kg/h). However, the bias is not statistically significant because the 95% confidence interval includes $x = 0$. The limits of agreement for aircraft and SOE are ± 65.4 kg/h (dash-dot lines in Figure B.3(a)), indicating that aircraft may predict a FLER 82.3 kg/h greater than or 48.5 kg/h less than SOE.

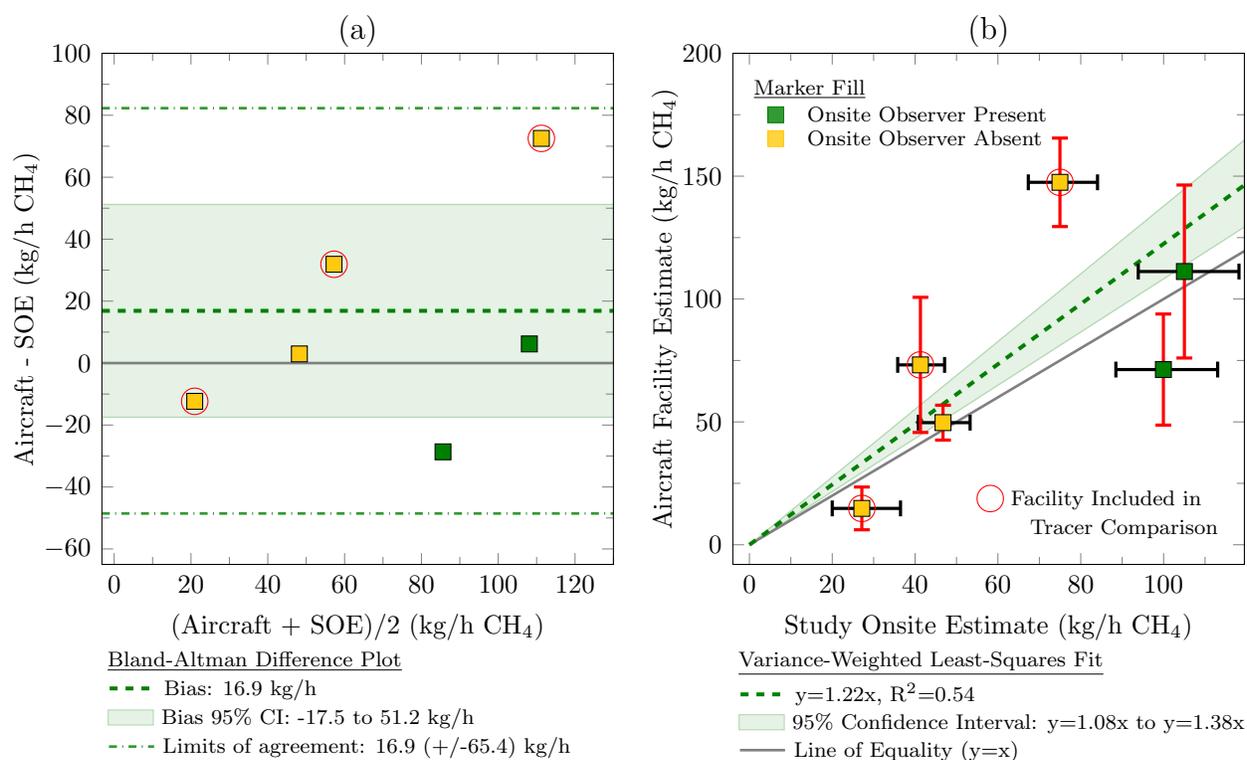


Figure B.3: Aircraft Facility Estimate vs Study Onsite Estimate Dehy In Difference Plot

In Figure B.3(b) a VWLS regression (dashed line) is performed on aircraft and SOE. The slope of the regression (aircraft = 1.22·SOE, $R^2 = 0.54$) is greater than unity, indicating that aircraft predicts higher FLER than FLER. The 95% confidence interval (shaded region) on the regression slope (aircraft = 1.08·SOE to aircraft = 1.38·SOE) does not include the line of equality ($y = x$), indicating that aircraft predicts higher FLER than SOE at the 95% confidence level.

Table B.2 shows the contribution of SOE component categories to the cumulative SOE for facilities included in the aircraft to SOE method comparison.

Table B.2: Emissions by category in the SOE for the 6 gathering station included in the AFE SOE method comparison.

CH ₄ Emission Source	Onsite Direct Measurements	Simulated Direct Measurements			Simulated
		Above Hi-Flow Range	Below Hi-Flow Range	Observed Not Measured	
Compressor Units	11.9	-	0.3	3.3	-
Pressure Relief Valves	0.6	-	-	-	-
Rod Packing Vents	17.8	12.3	0.1	-	-
Dehydrator Units	2.2	-	0.2	-	-
Other	0.6	-	0.0	-	-
Pig Launchers/Receivers	-	-	-	-	-
Piping or Gas Lines	4.6	-	0.2	-	-
Separators	2.9	-	0.2	-	-
Tanks	3.9	-	0.0	-	-
Combustion Slip	-	-	-	-	253.1
Crankcase Vents	-	-	-	-	27.8
Dehydrator Regenerator Vents	-	-	-	-	53.2
Cumulative Study Onsite Estimate	44.5	12.3	0.9	3.3	395.1 kg/h
Cumulative Aircraft Facility Estimate					467.7 kg/h

B.2 VARIANCE-WEIGHTED LEAST-SQUARES REGRESSION

The variance-weighted least-squares (VWLS) regression used in method comparisons employs the method described in Neri et al.⁵⁴, and summarized here for convenience. Briefly, the sum of the squared orthogonal distances between each data point $P(x_i, y_i)$ and the line of best fit $y = ax + b$ (i.e. the VWLS fit) is minimized, accounting for the uncertainty δ (standard deviation) in both x and y data, $(\delta x_i, \delta y_i)$, by weighting each data point $P(x_i, y_i)$ by W_i . W_i is defined as the squared inverse of the orthogonal distance between the line of best fit and the data point, d_i :

$$W_i = \frac{1}{(\delta d_i)^2} \quad (\text{B.1})$$

Where d_i is given by:

$$d_i = \frac{(ax_i - y_i + b)}{\sqrt{a^2 + 1}} \quad (\text{B.2})$$

As illustrated in Figure B.4.

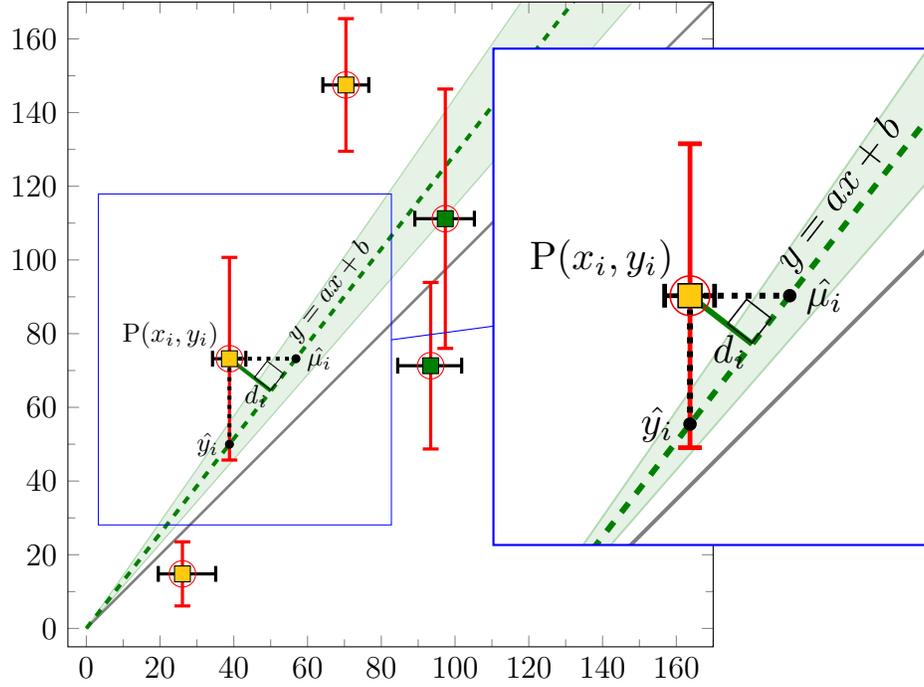


Figure B.4: Example data point illustrating the variance-weighted least-squares method.

By applying the propagation of error law to d_i :

$$\delta d_i = \frac{\partial d_i}{\partial x_i} \delta x_i + \frac{\partial d_i}{\partial y_i} \delta y_i = \frac{a}{\sqrt{a^2 + 1}} \delta x_i + \frac{1}{\sqrt{a^2 + 1}} \delta y_i \quad (\text{B.3})$$

and assuming independent and random error; the error terms are added in quadrature to avoid an overestimate of the overall uncertainty:

$$(\delta d_i)^2 = \frac{a^2}{a^2 + 1} (\delta x_i)^2 + \frac{1}{a^2 + 1} (\delta y_i)^2 \quad (\text{B.4})$$

The FLER regression then becomes an exercise in minimizing F :

$$F = \sum_i^N \left(\frac{ax_i - y_i + b}{\sqrt{a^2 + 1}} \right)^2 \quad (\text{B.5})$$

where each of the N experimental data points $P(x_i, y_i)$ are weighted by:

$$W_i = \frac{a^2 + 1}{a^2(\delta x_i)^2 + (\delta y_i)^2} \quad (\text{B.6})$$

The minimization is carried out using the bisection method outlined in Press et al.⁹⁷ The minimization routine was implemented in C[‡], and was compared to the test case provided in Neri et al. The comparison indicated that the minimization routine was implemented successfully.

Table B.3: Variance-weighted least-squares regression minimization routine testing results, indicating successful test data reproduction.

Calculated value	Neri et al.	Present Paper	Exact Solution
a	-0.480 553 402 6	-0.480 533 407 446 273 49	-0.480 533 407
b	5.479 910 219 48	5.479 910 22 403 321 36	5.479 910 22

Additionally, we define a “total coefficient of determination” R^2 as:

$$R^2 = 1 - \frac{SSE}{SST} \quad (\text{B.7})$$

Where SSE and SST include both x and y errors by defining:

$$SST = \sum (y_i - \bar{y})^2 + \sum (x_i - \bar{x})^2 \quad (\text{B.8})$$

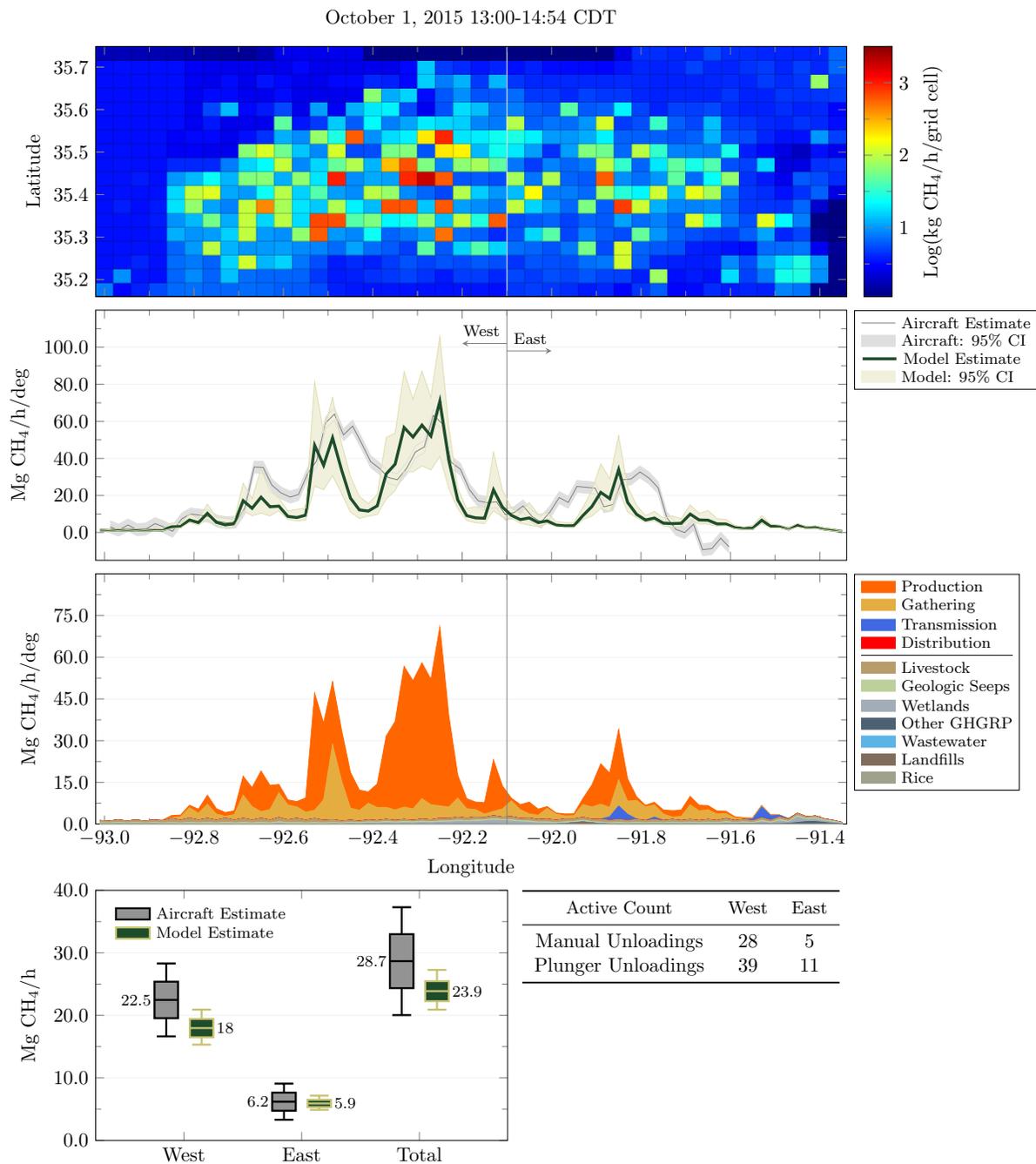
$$SSE = \sum (y_i - \hat{y}_i)^2 + \sum (x_i - \hat{\mu}_i)^2 \quad (\text{B.9})$$

APPENDIX C

RECONCILING TOP-DOWN AND BOTTOM-UP METHANE EMISSION ESTIMATES FROM NATURAL GAS OPERATIONS IN THE FAYETTEVILLE SHALE

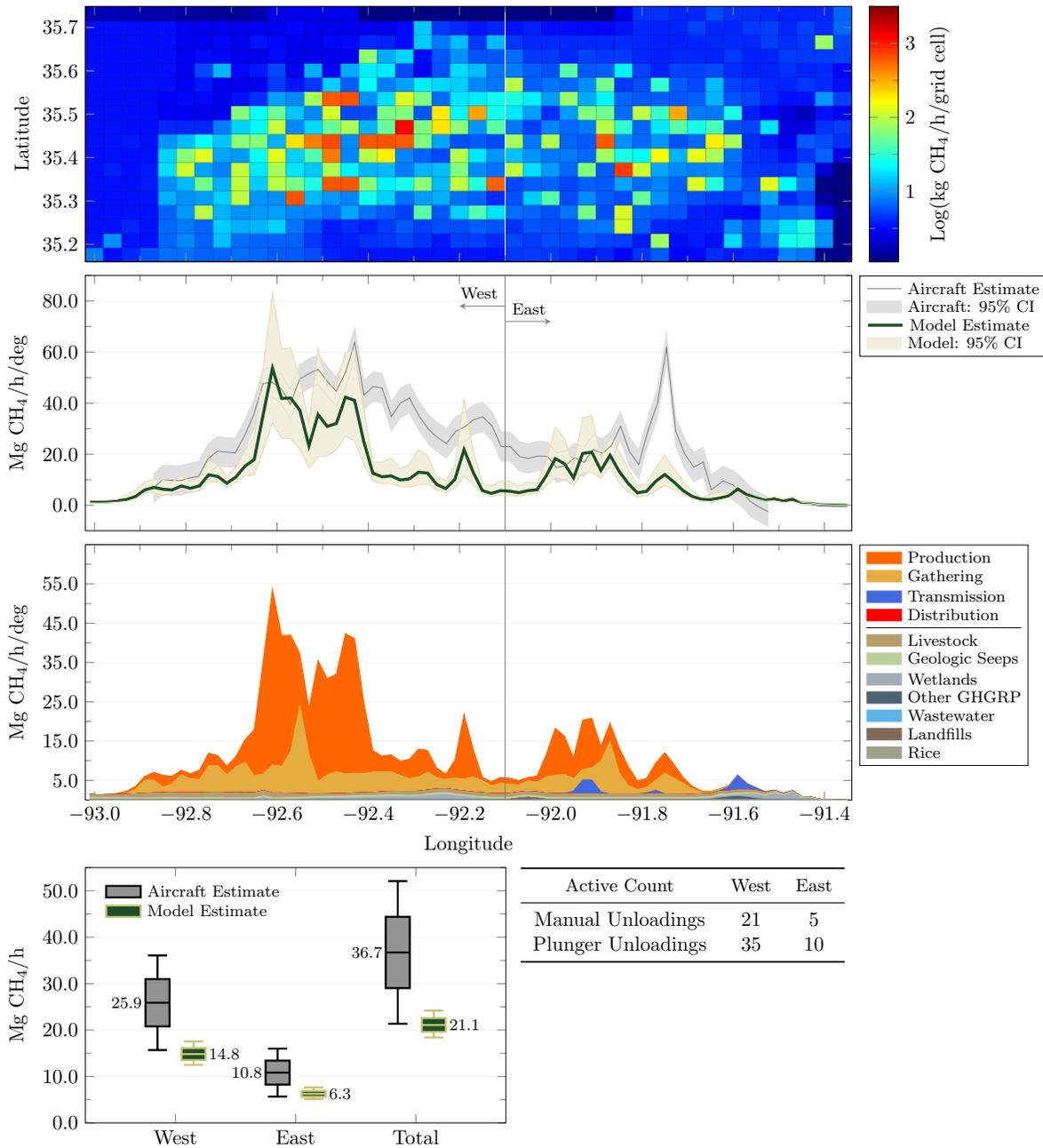
C.1 GROUND-LEVEL AREA ESTIMATE (GLAE) RESULTS CORRESPONDING AIRCRAFT MASS BALANCE FLIGHTS

C.1.1 FLIGHT WINDOW GLAE RESULTS: OCTOBER 1ST, 2015



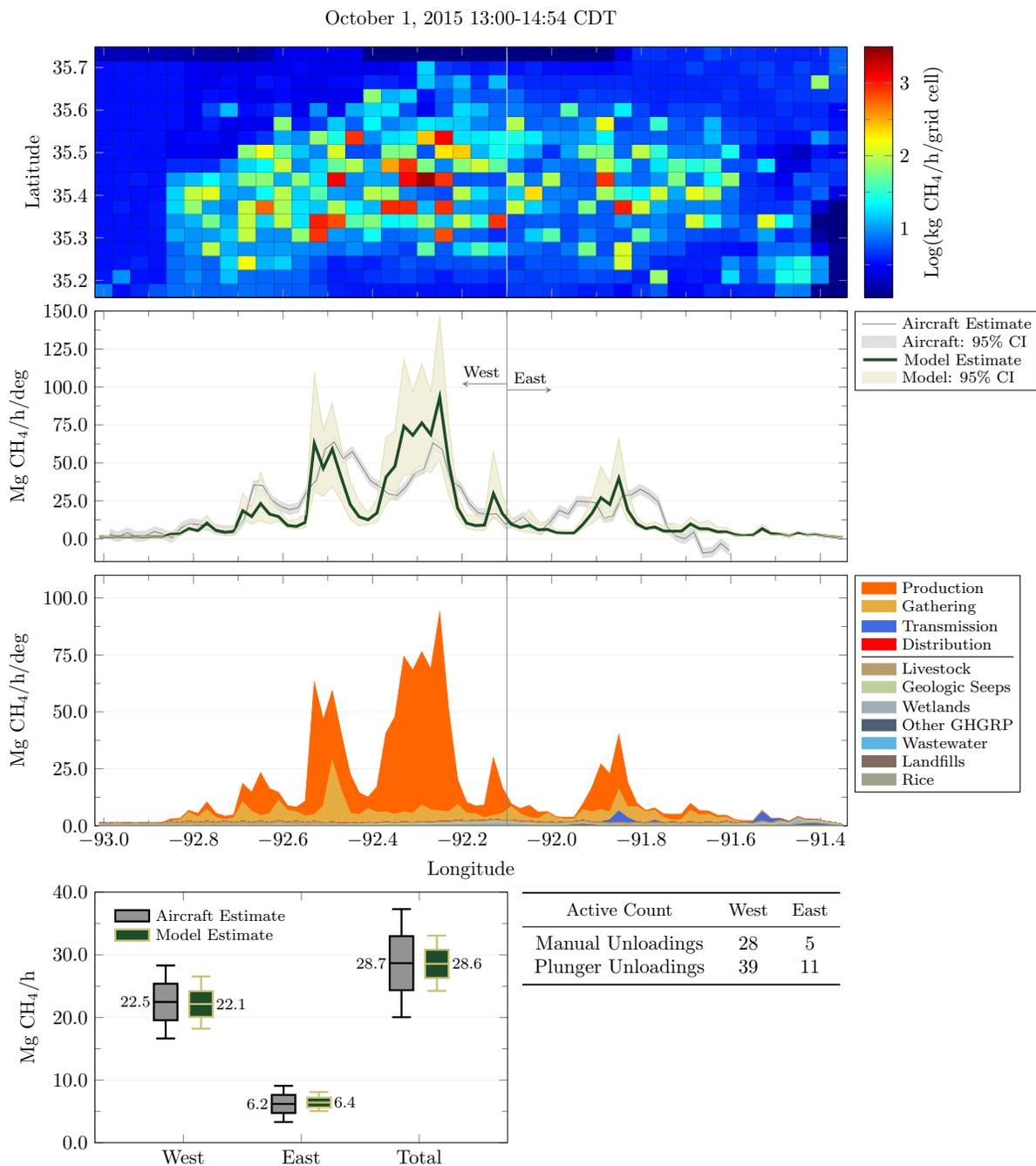
C.1.2 FLIGHT WINDOW GLAE RESULTS: OCTOBER 2ND, 2015

October 2, 2015 14:48-16:30 CDT



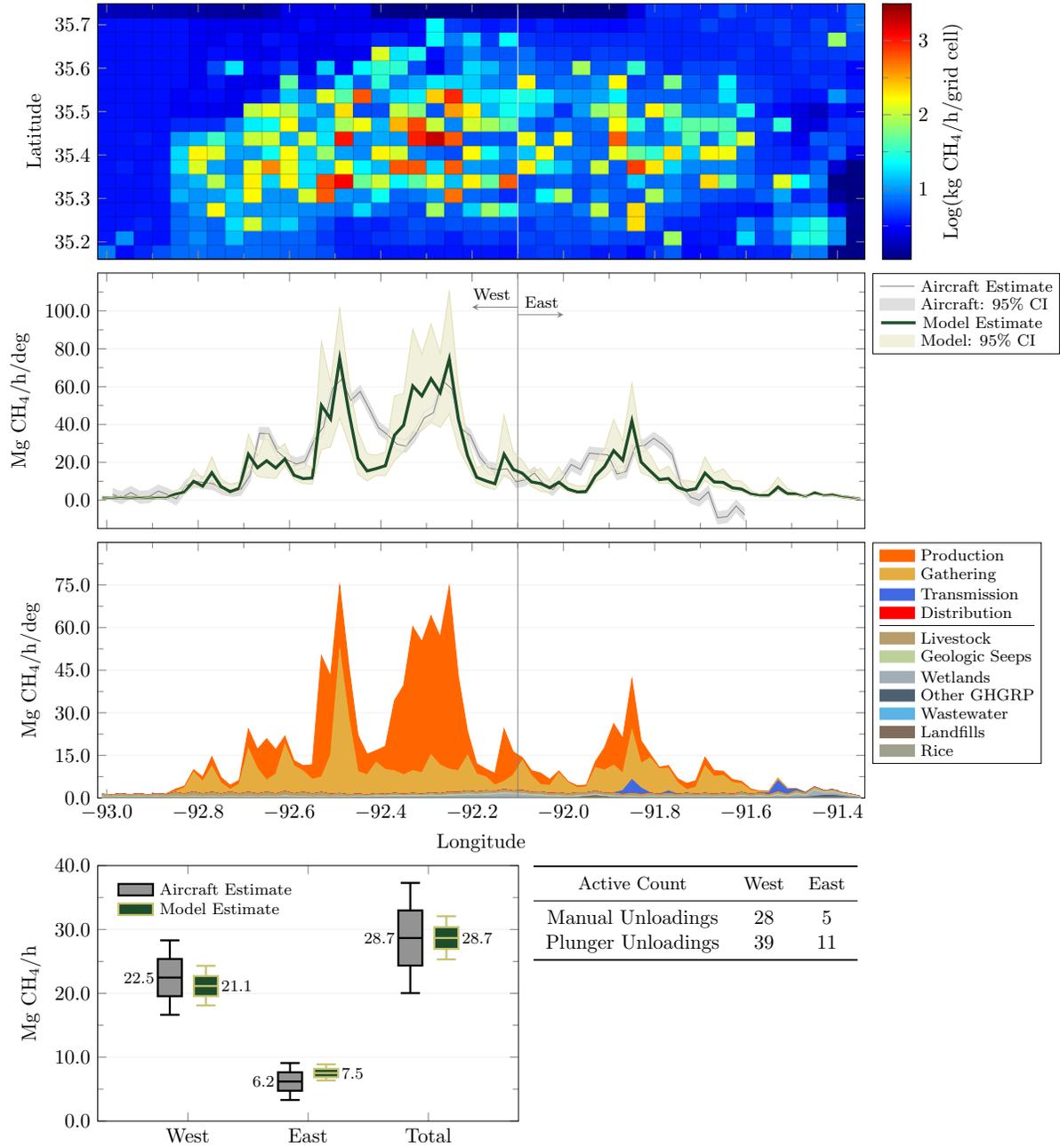
C.2 GROUND-LEVEL AREA ESTIMATE (GLAE) SENSITIVITY STUDIES CORRESPONDING TO THE OCTOBER 1ST, 2015 AIRCRAFT MASS BALANCE FLIGHT

C.2.1 SCENARIO 1: MANUAL LIQUID UNLOADING EMISSION RATES INCREASED 37%



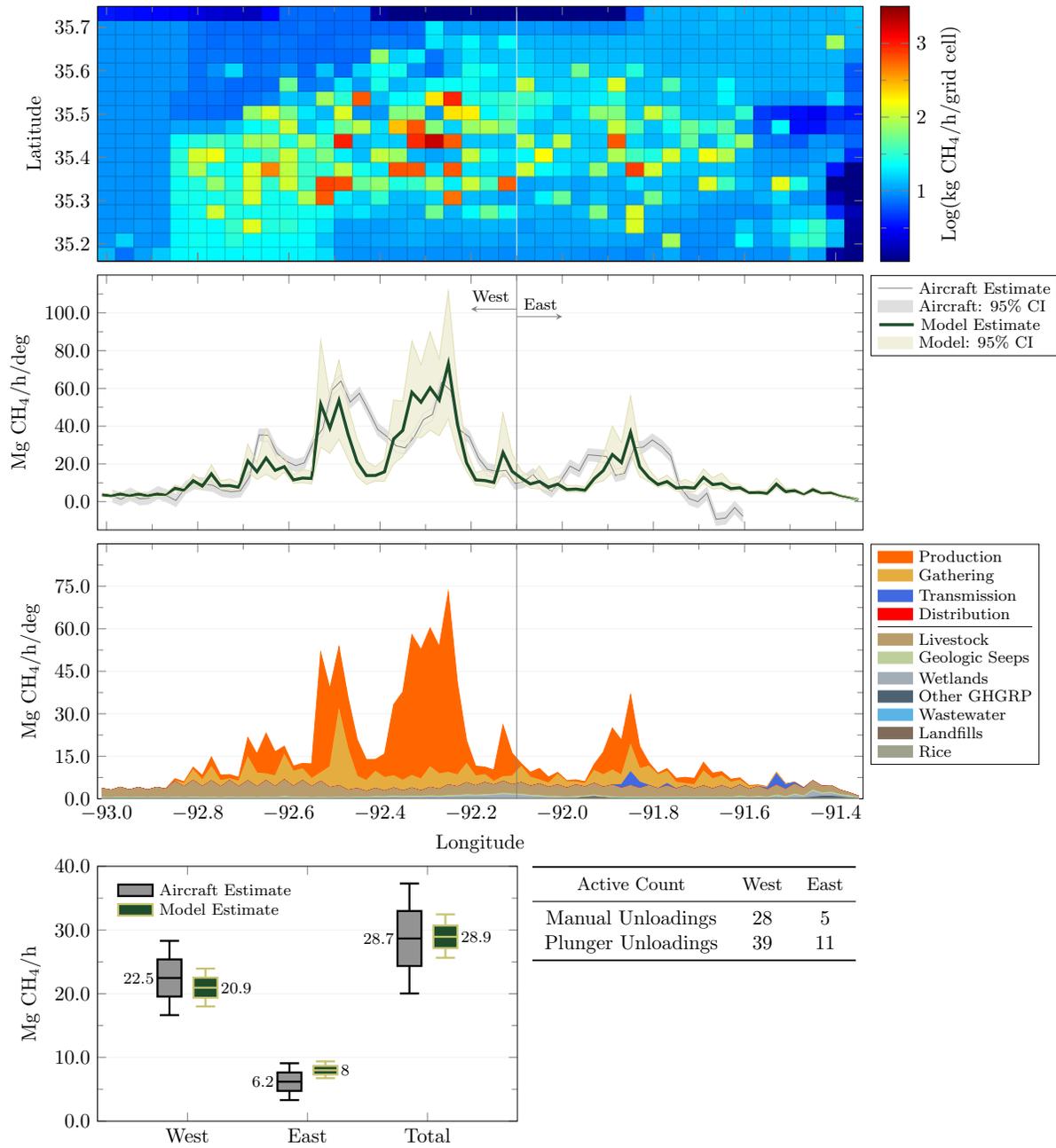
C.2.2 SCENARIO 2: GATHERING STATION EMISSION RATES INCREASED 89%

October 1, 2015 13:00-14:54 CDT

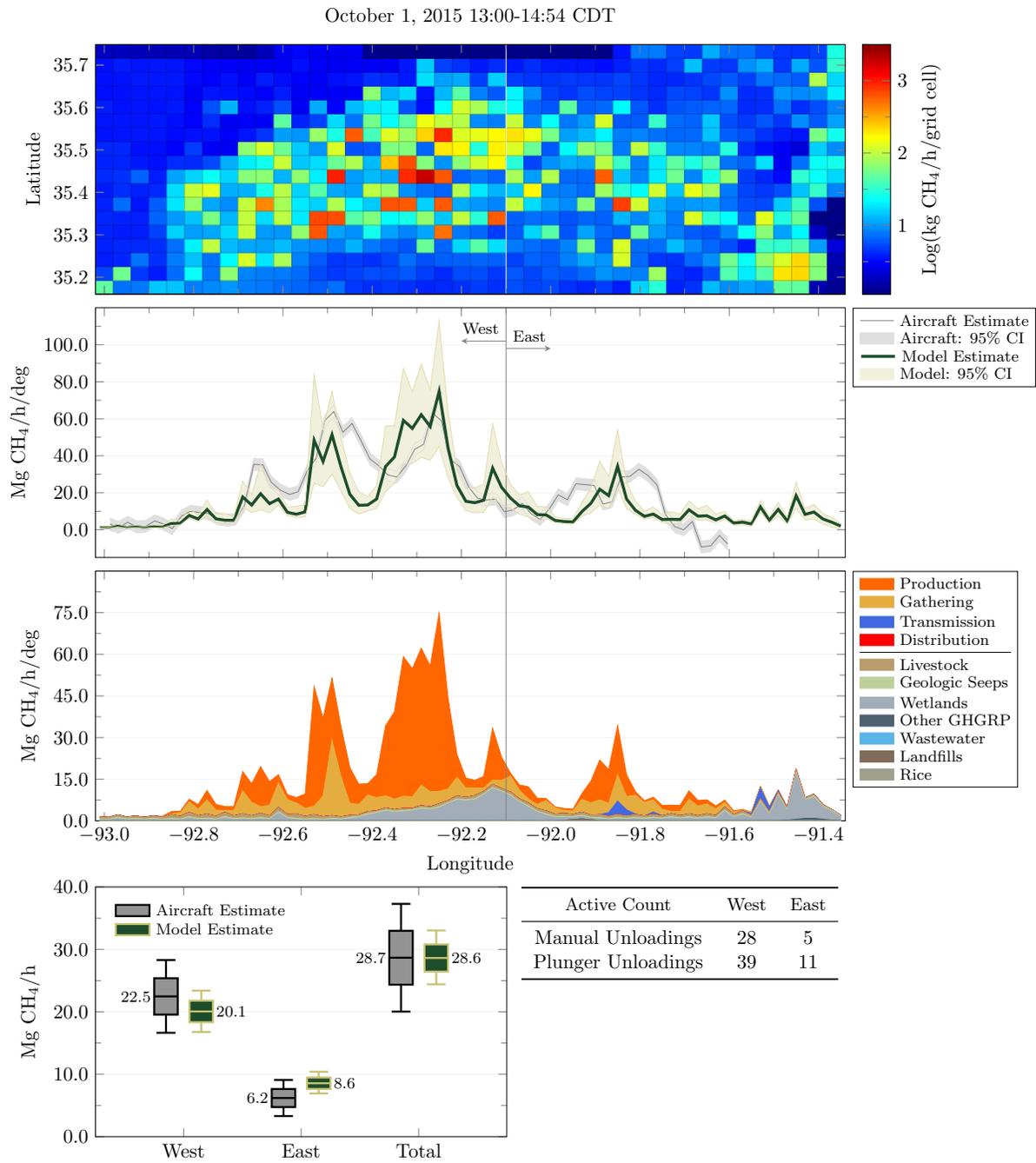


C.2.3 SCENARIO 3: LIVESTOCK EMISSION RATES INCREASED 4X

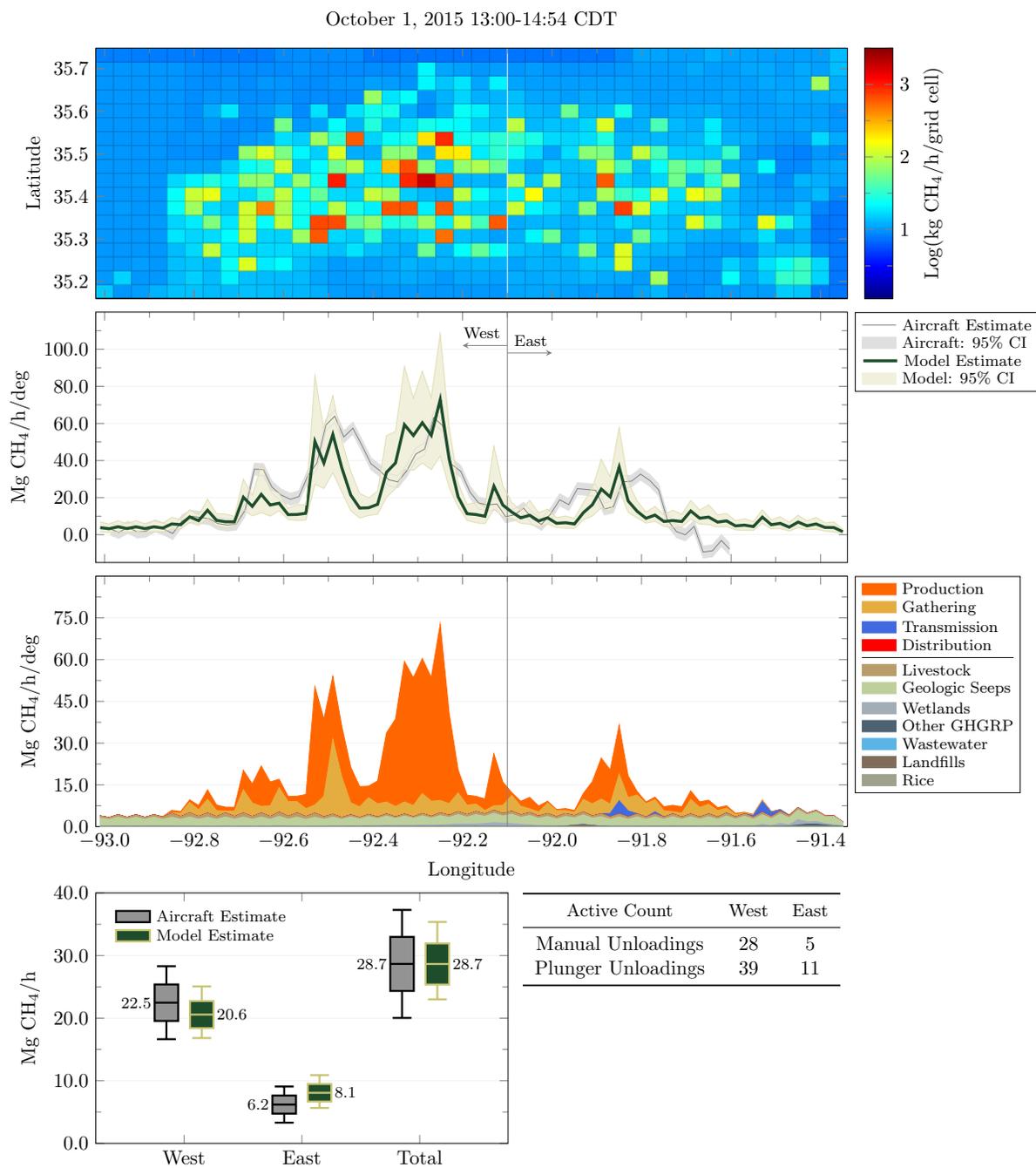
October 1, 2015 13:00-14:54 CDT



C.2.4 SCENARIO 4: WETLAND EMISSION RATES INCREASED 8.5X

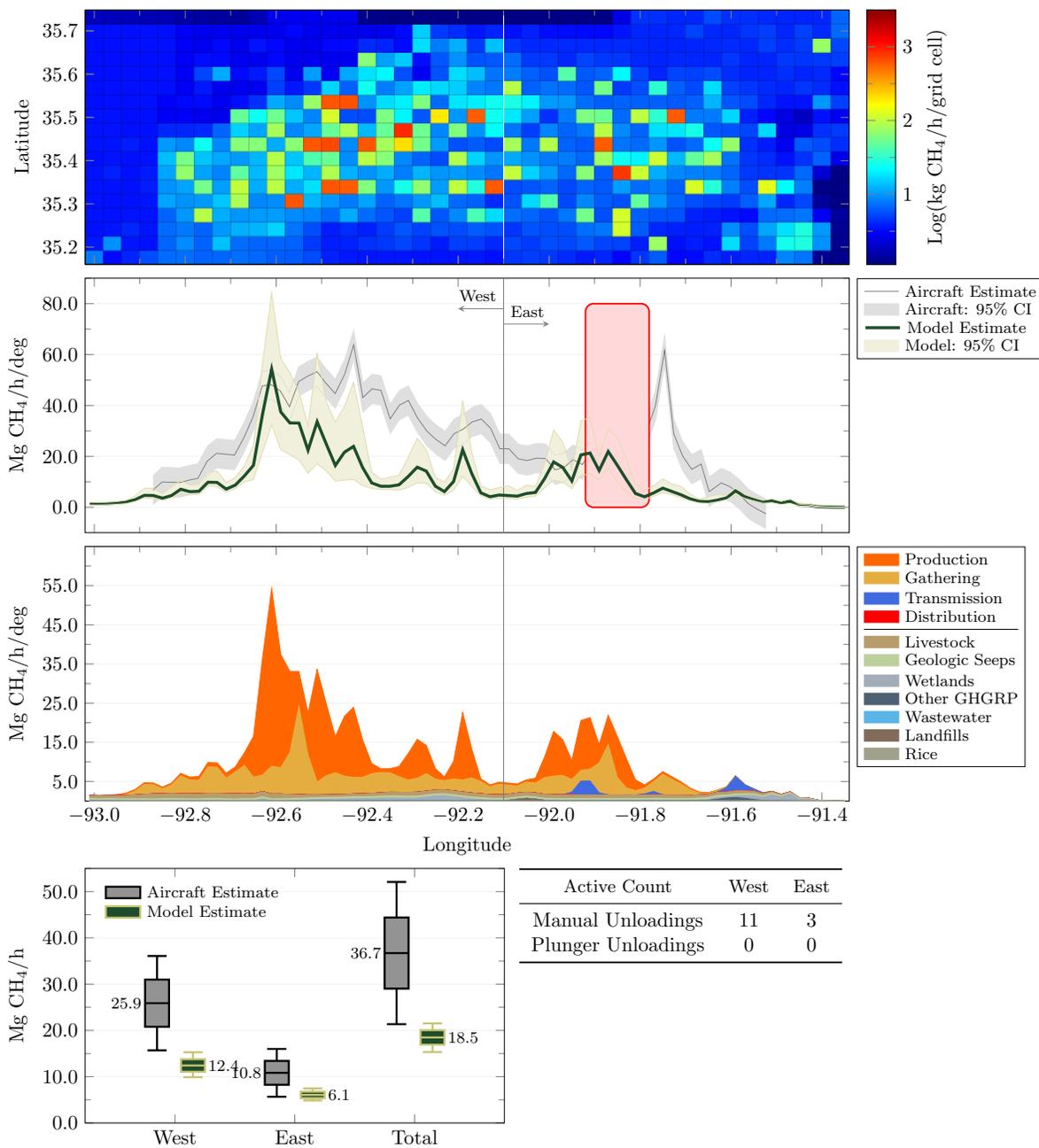


C.2.5 SCENARIO 5: GEOLOGIC SEEP EMISSION RATES INCREASED 6.5X

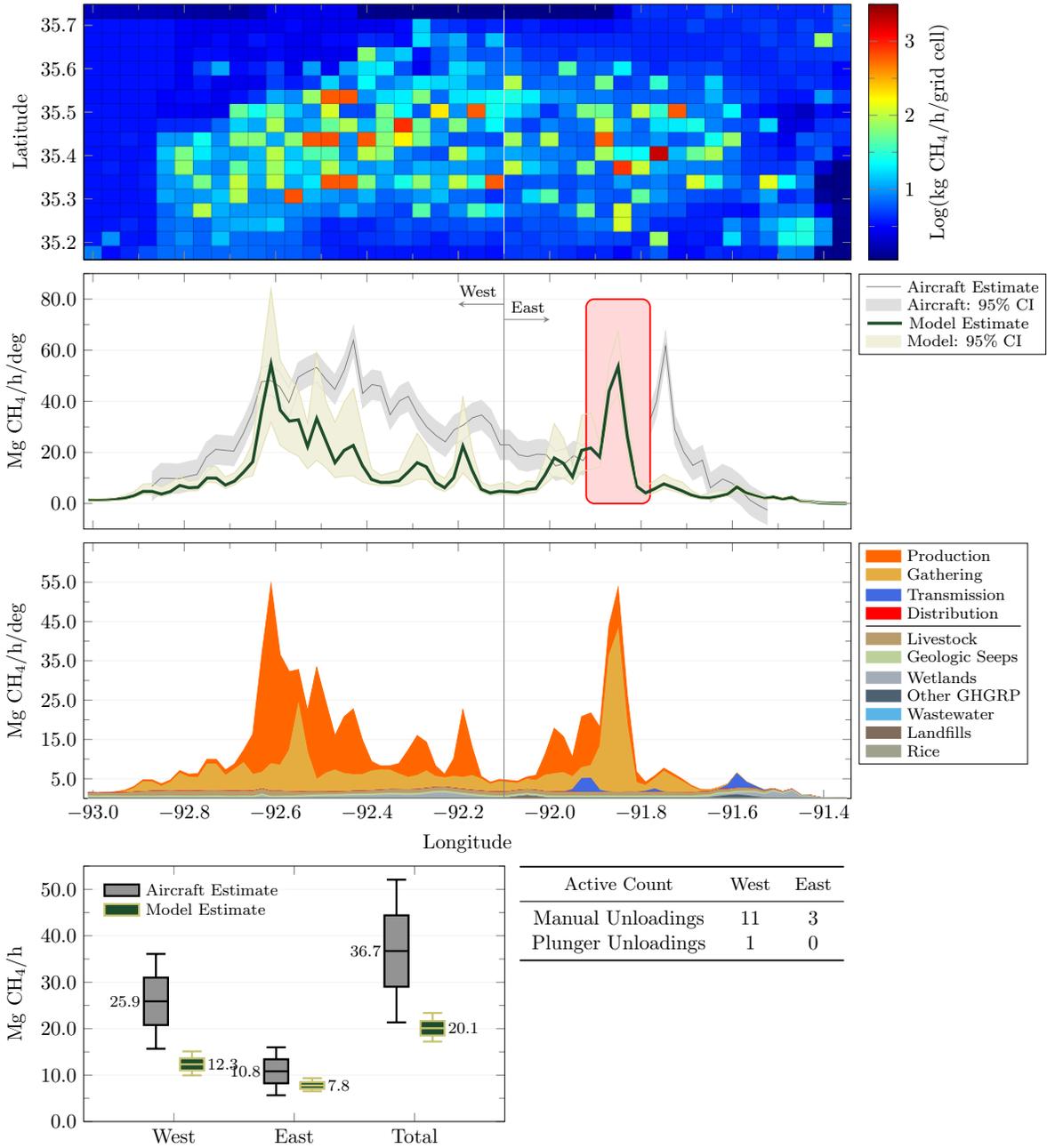


C.2.6 SCENARIO 6: TIMING OF SHORT-DURATION, HIGH-EMISSION RATE EVENTS

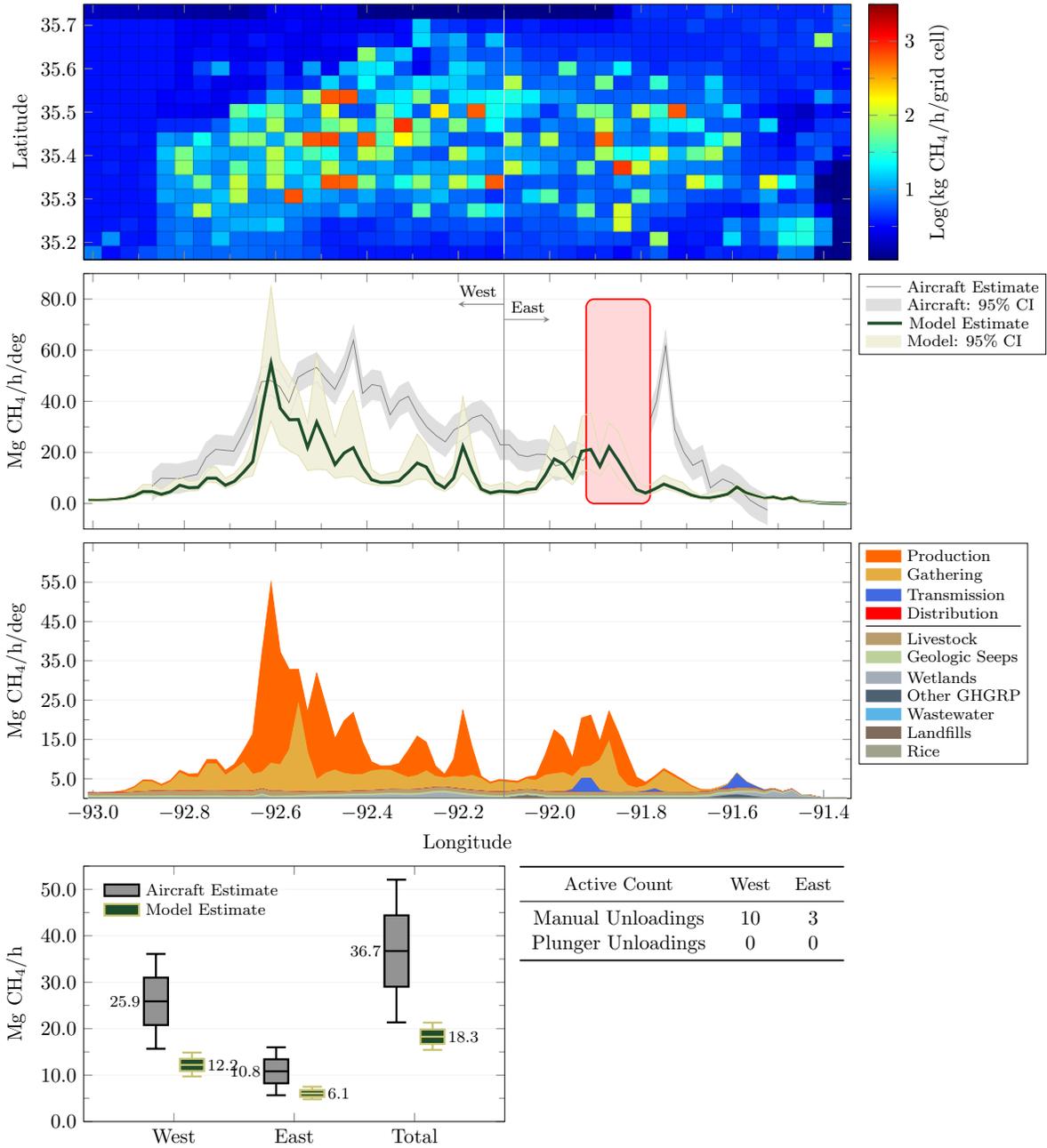
October 2, 2015 16:25-16:26 CDT



October 2, 2015 16:26-16:29 CDT

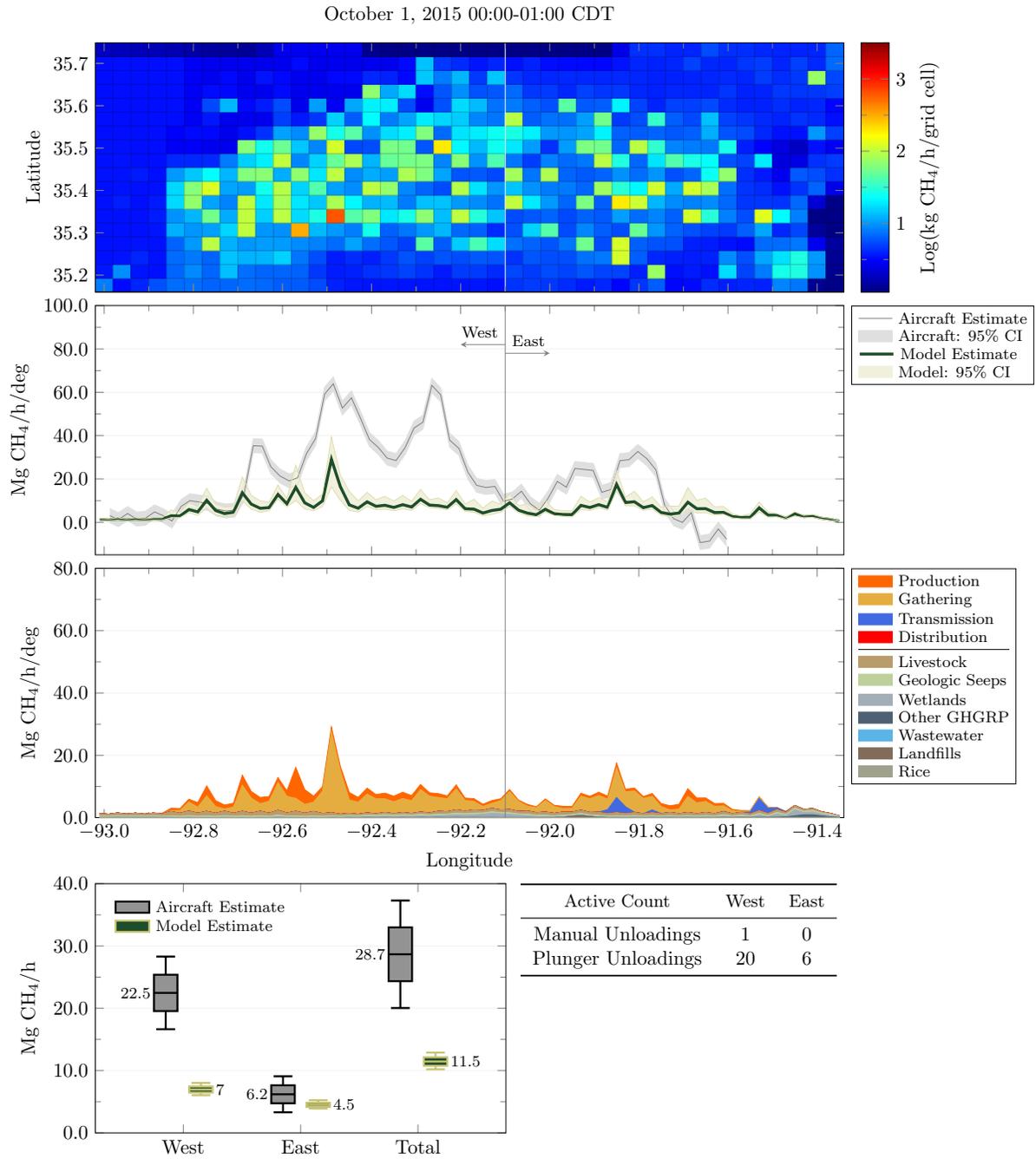


October 2, 2015 16:30-16:31 CDT

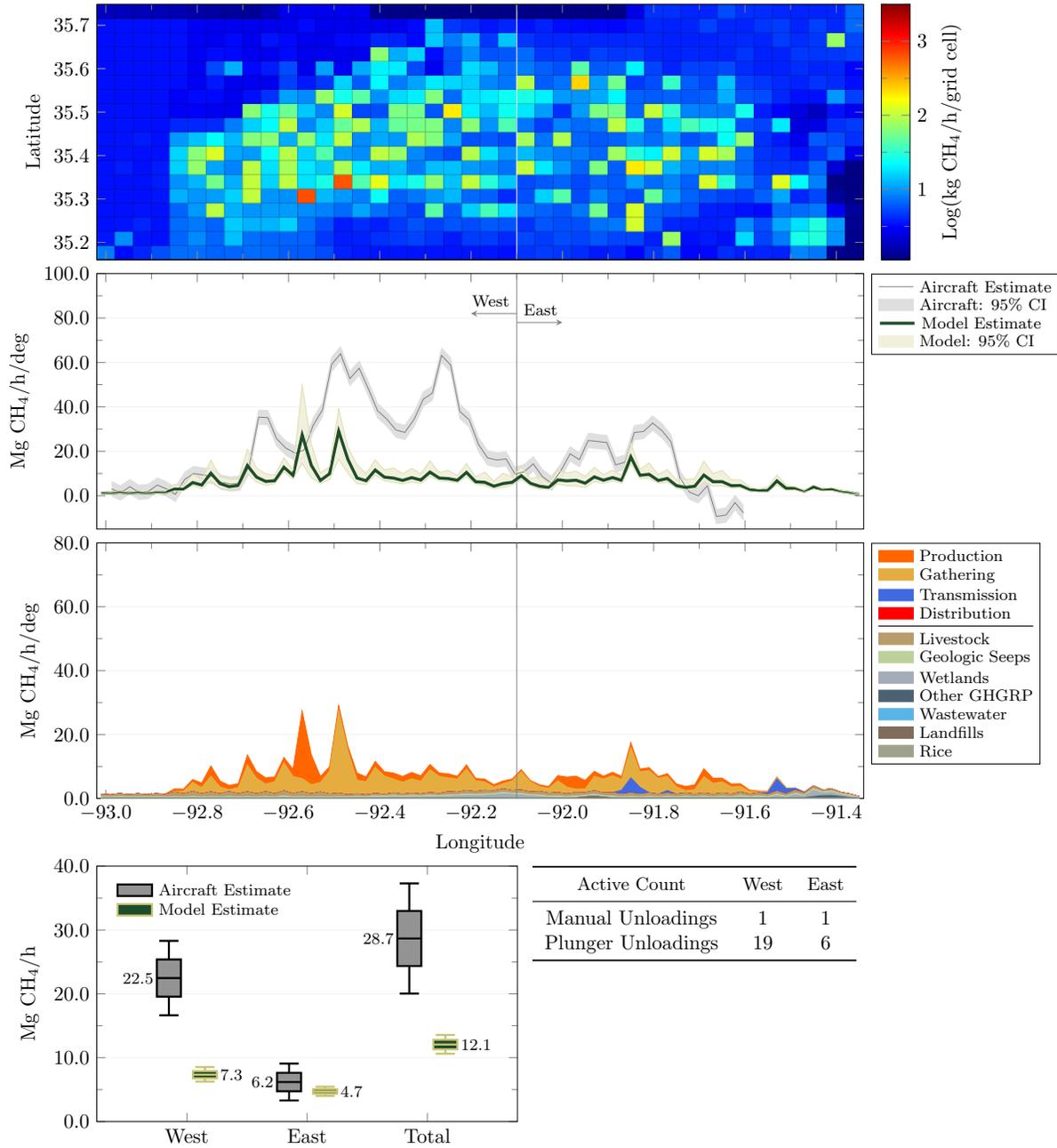


C.3 HOURLY GROUND-LEVEL AREA ESTIMATE (GLAE) RESULTS FOR OCTOBER 1ST AND 2ND, 2015

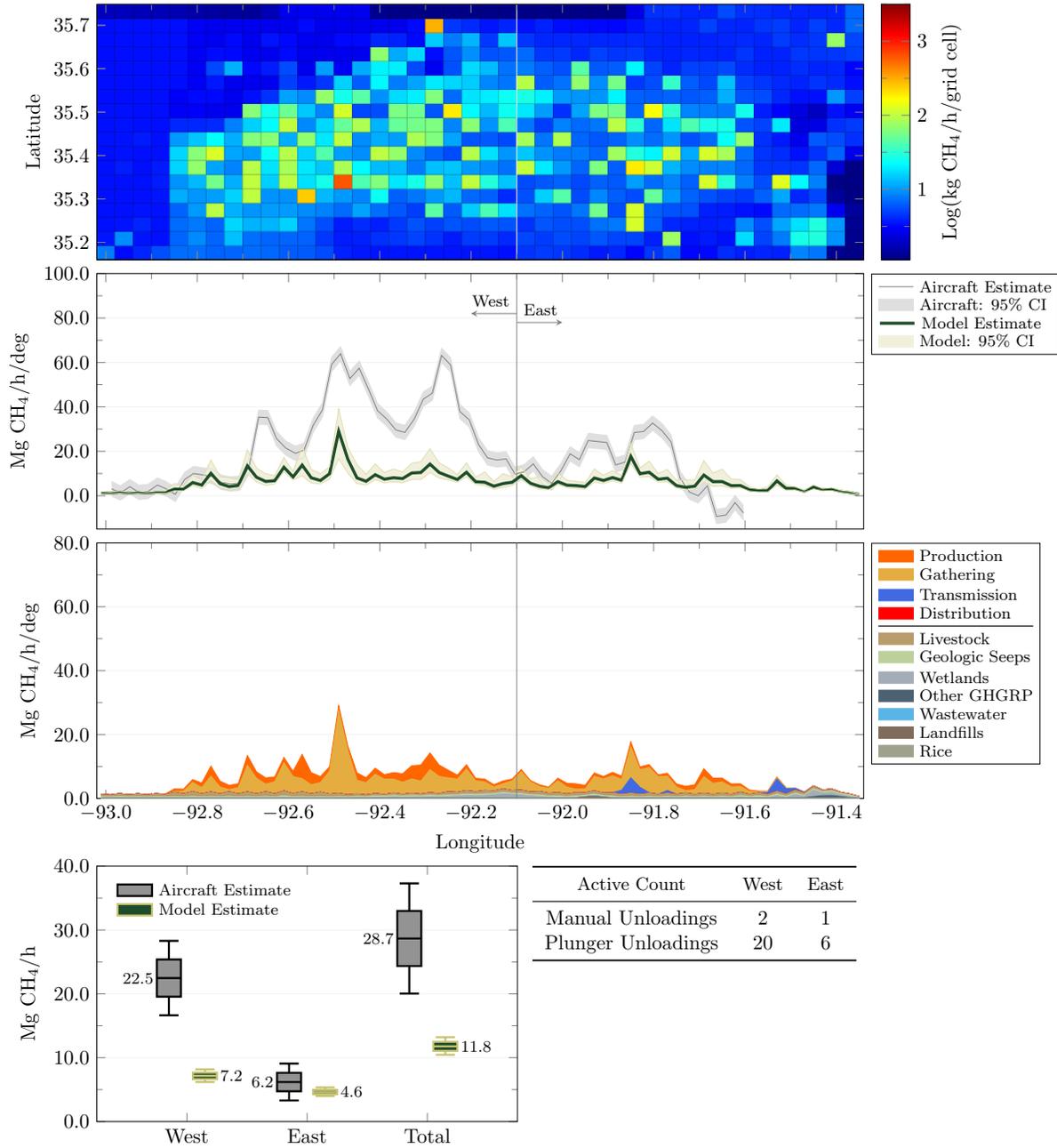
C.3.1 HOURLY GLAE RESULTS: OCTOBER 1ST, 2015



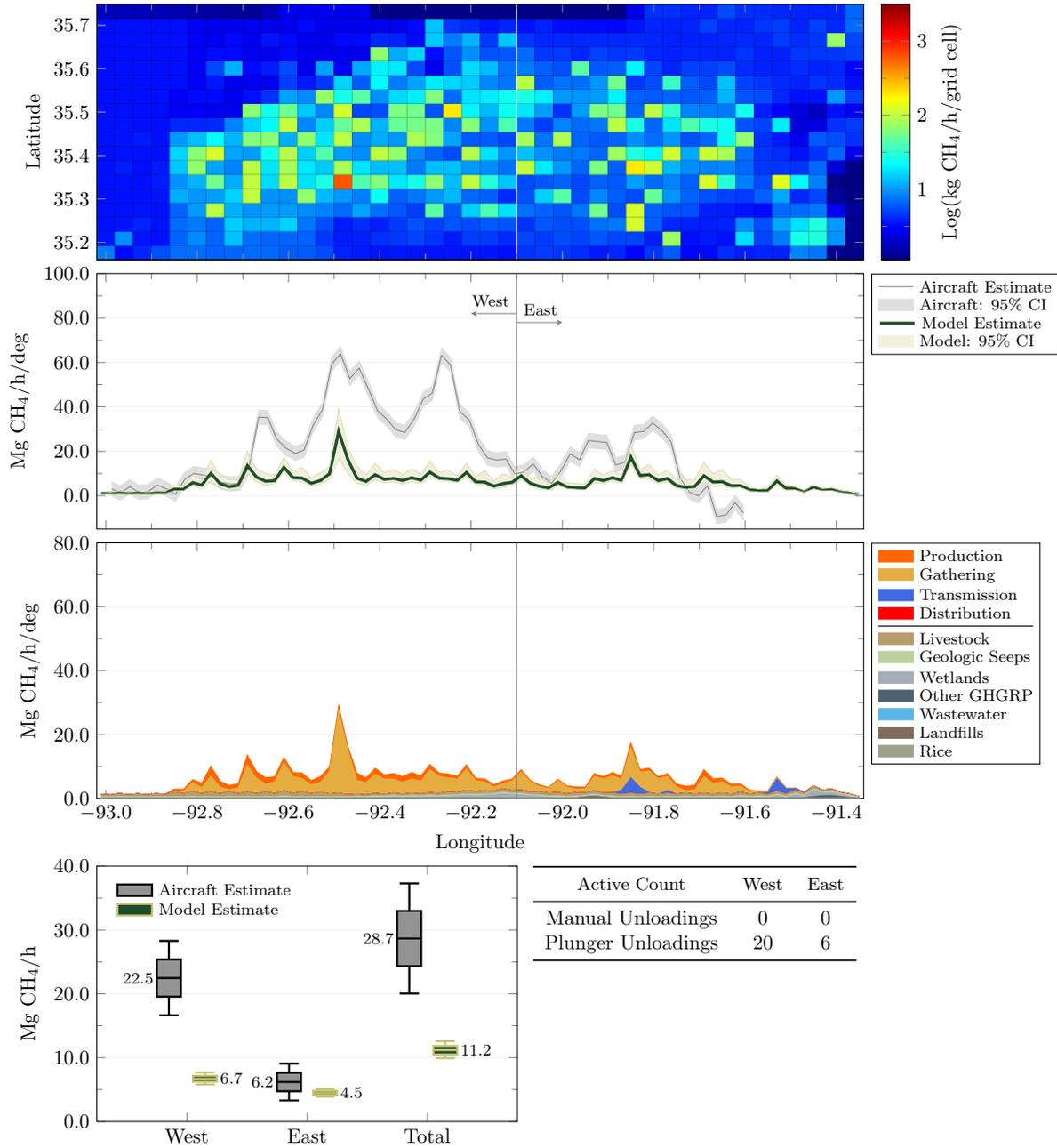
October 1, 2015 01:00-02:00 CDT



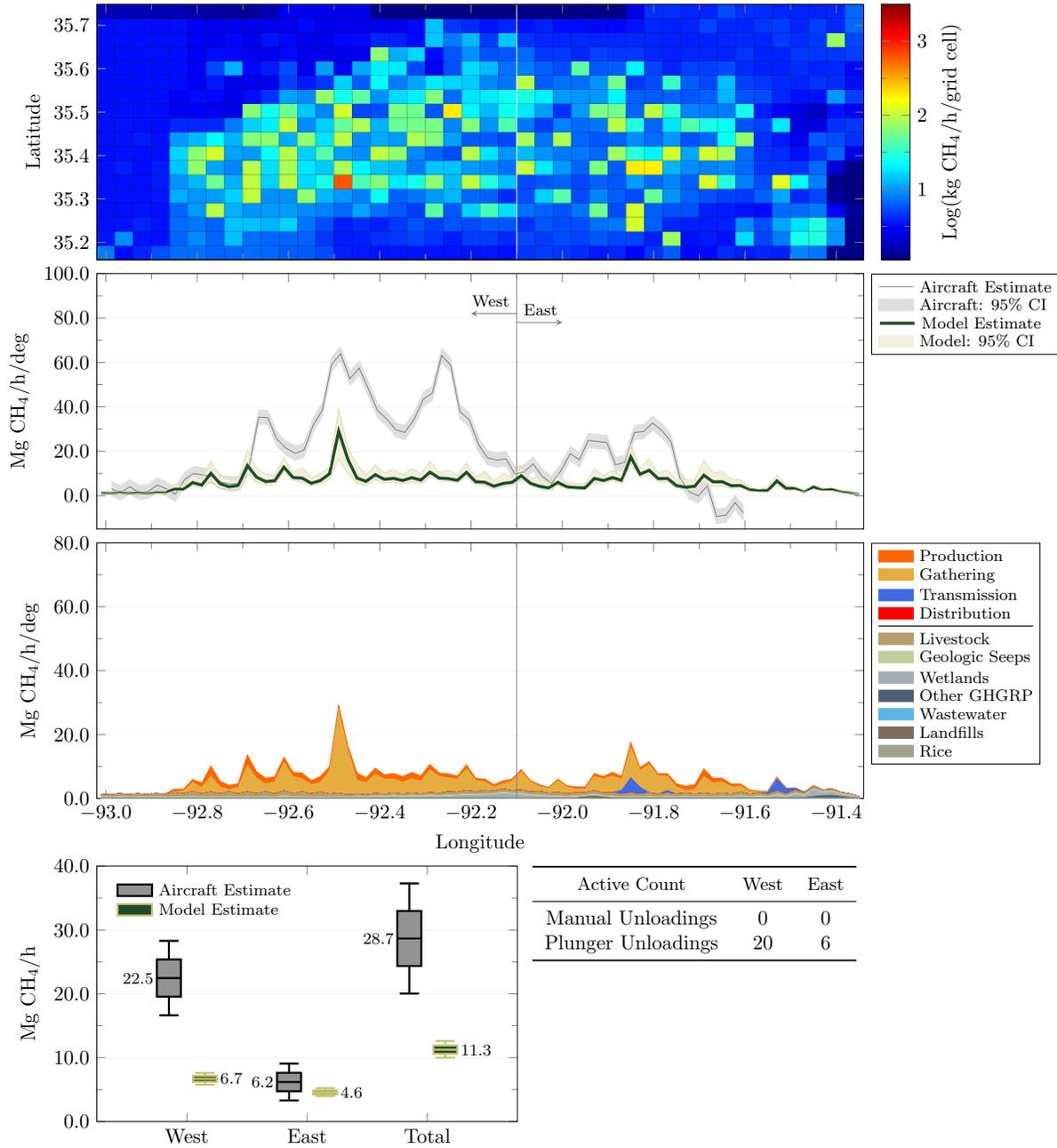
October 1, 2015 02:00-03:00 CDT



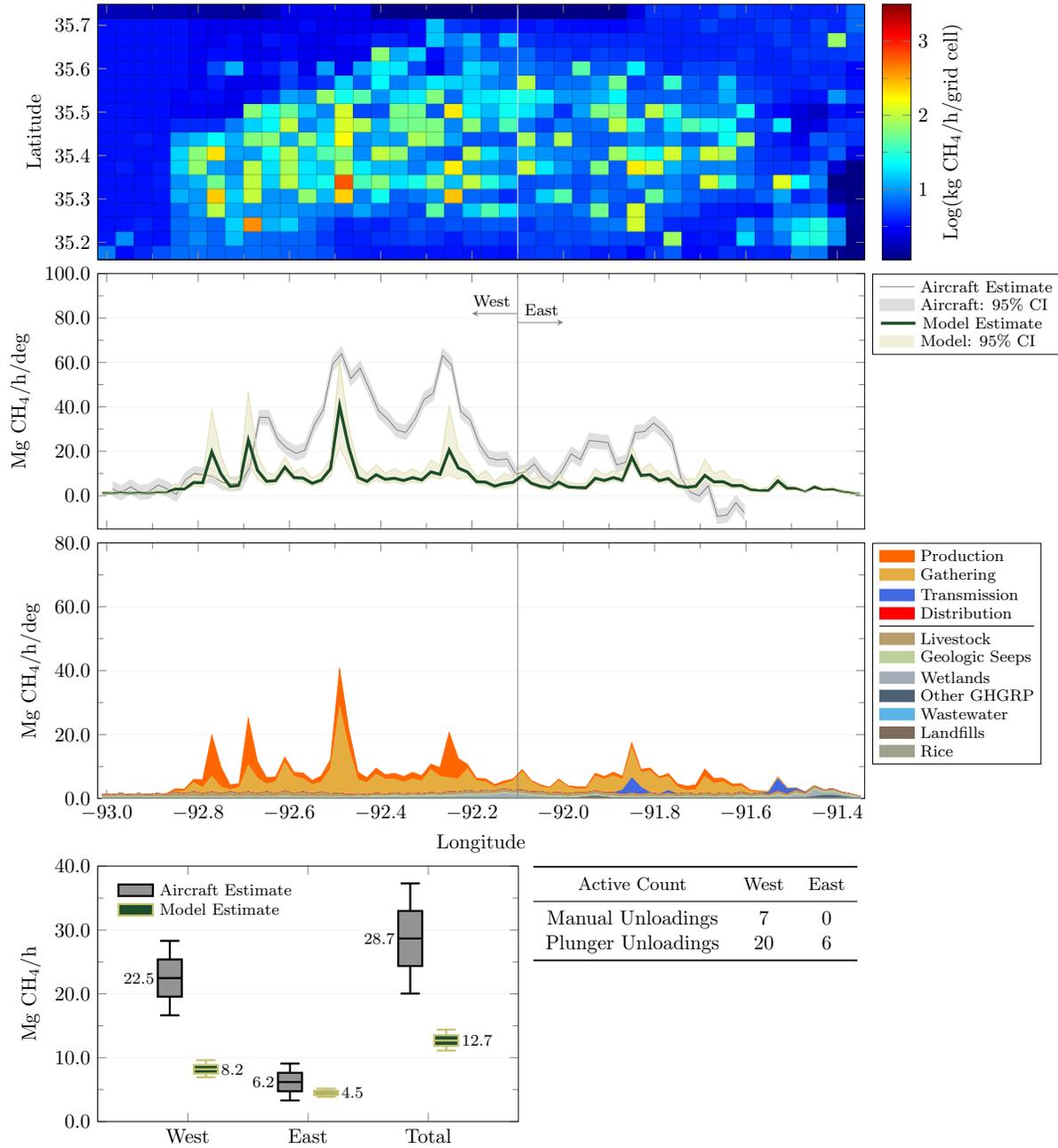
October 1, 2015 03:00-04:00 CDT



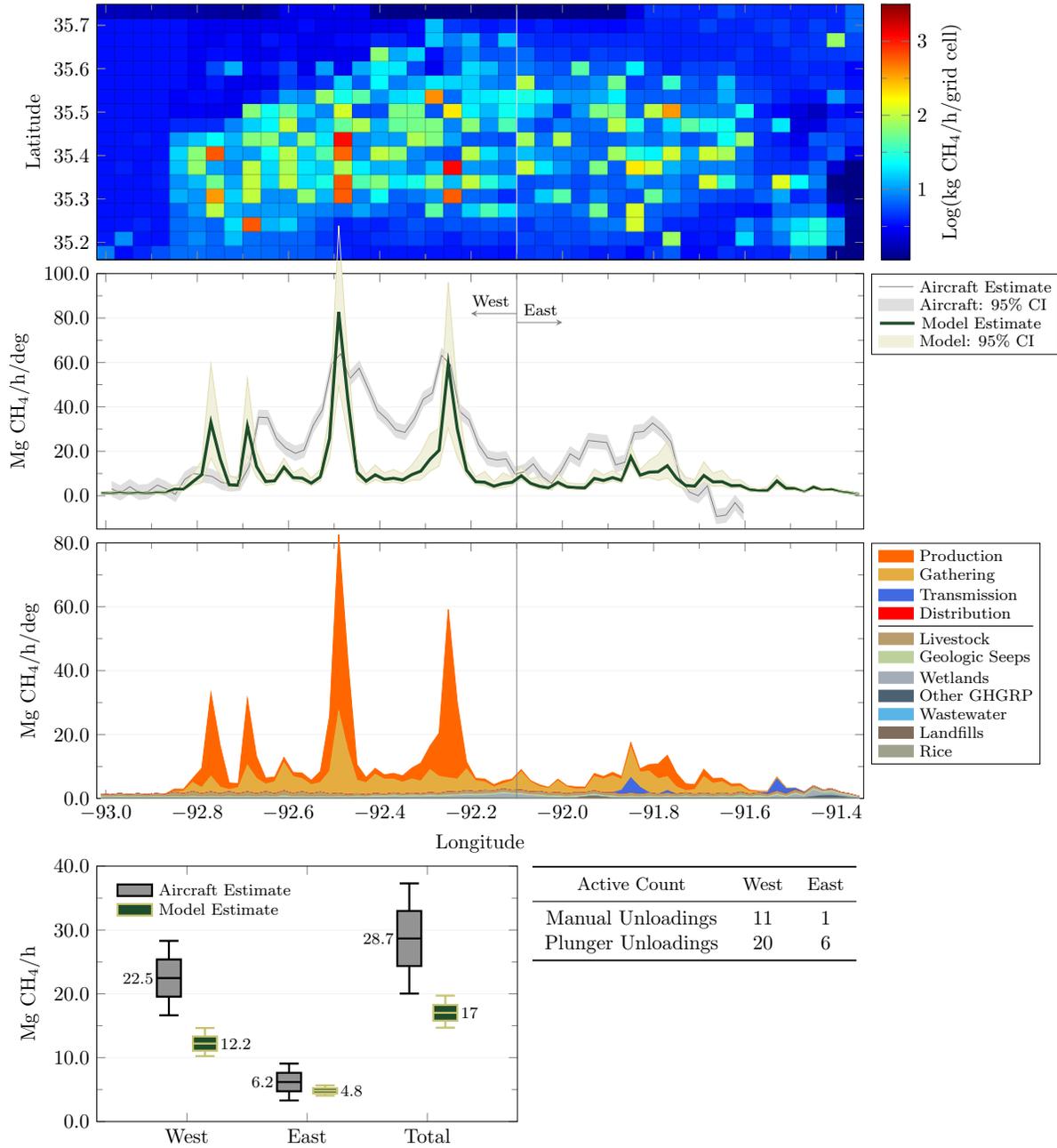
October 1, 2015 04:00-05:00 CDT



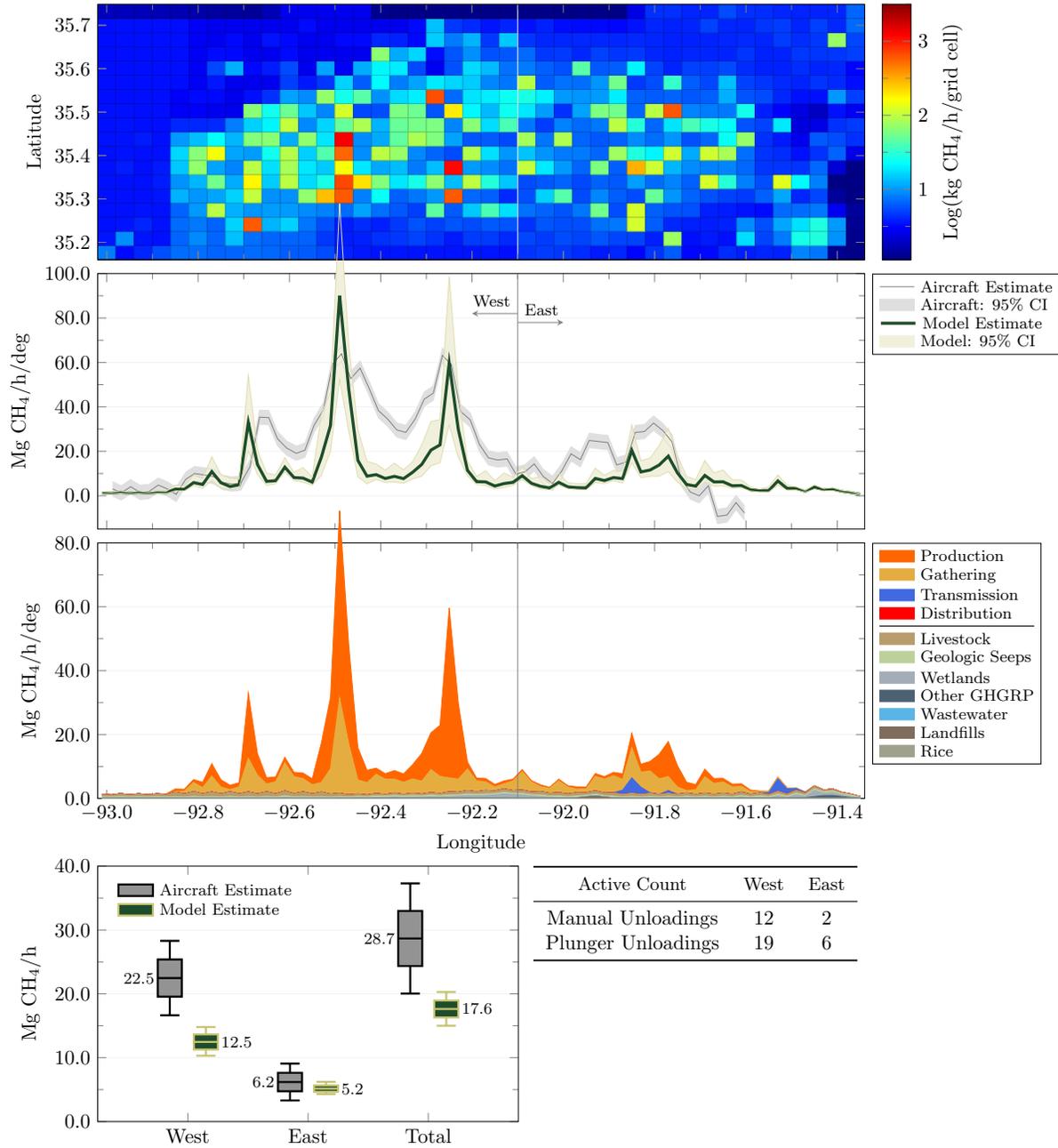
October 1, 2015 05:00-06:00 CDT



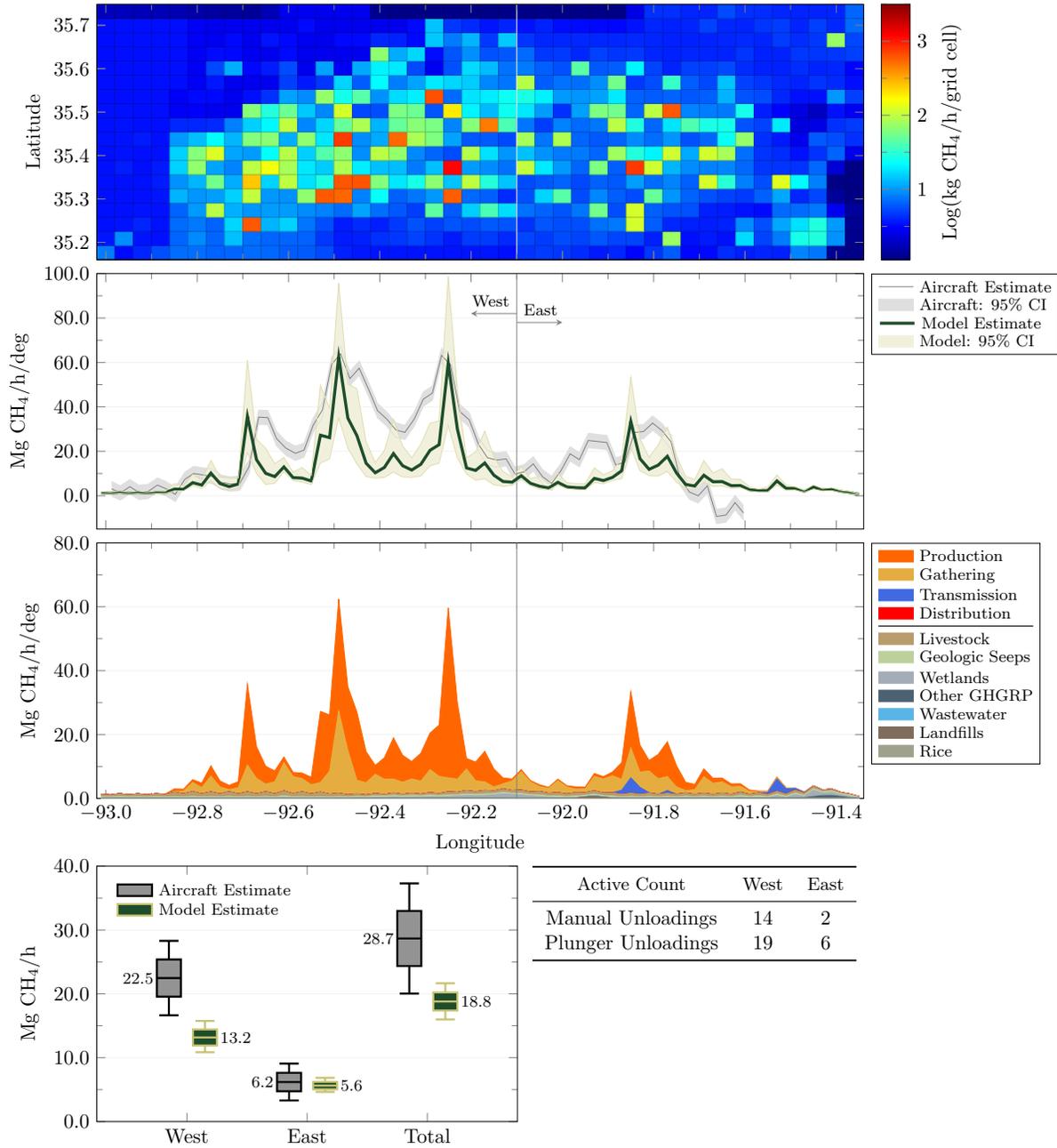
October 1, 2015 06:00-07:00 CDT



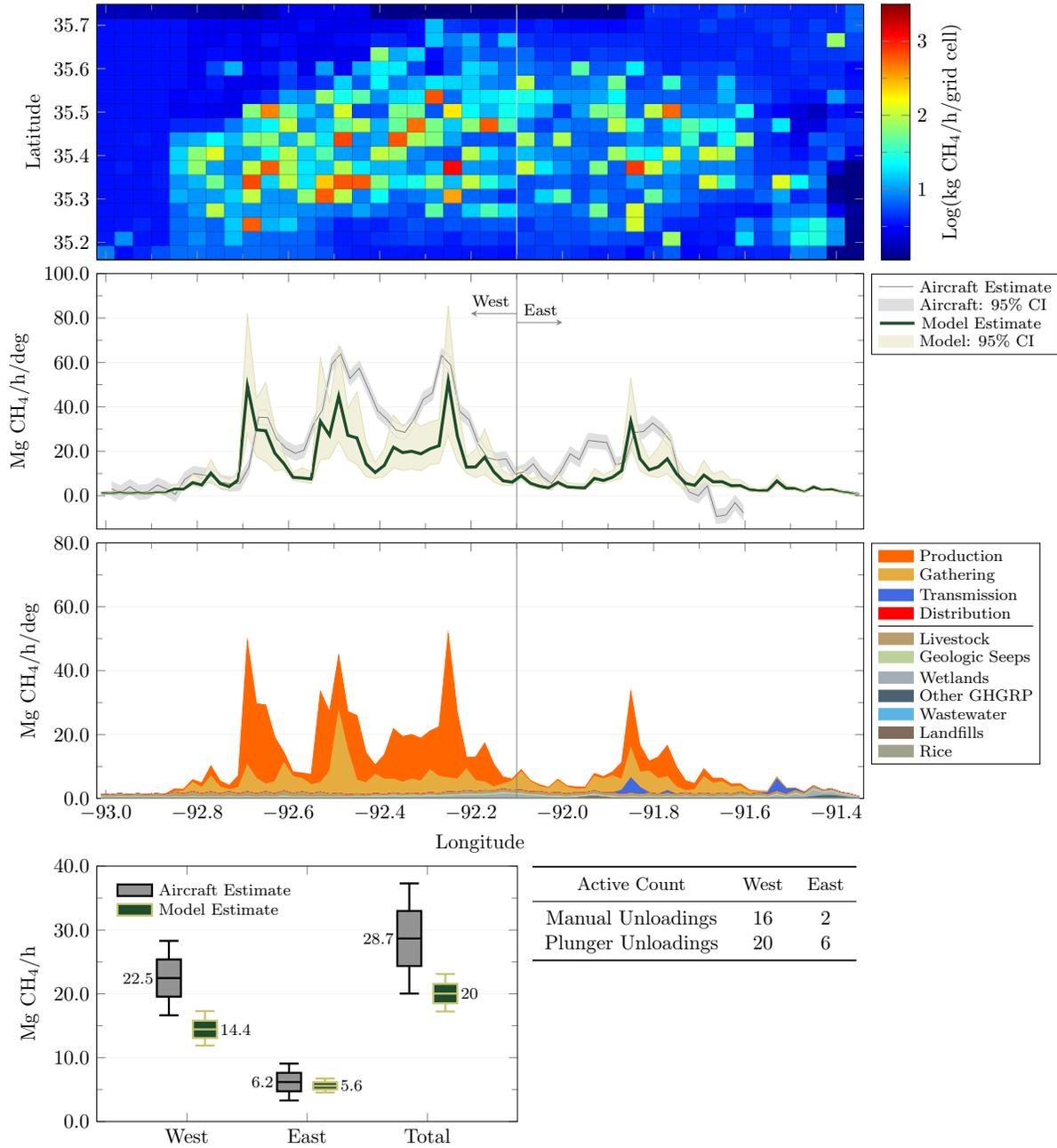
October 1, 2015 07:00-08:00 CDT



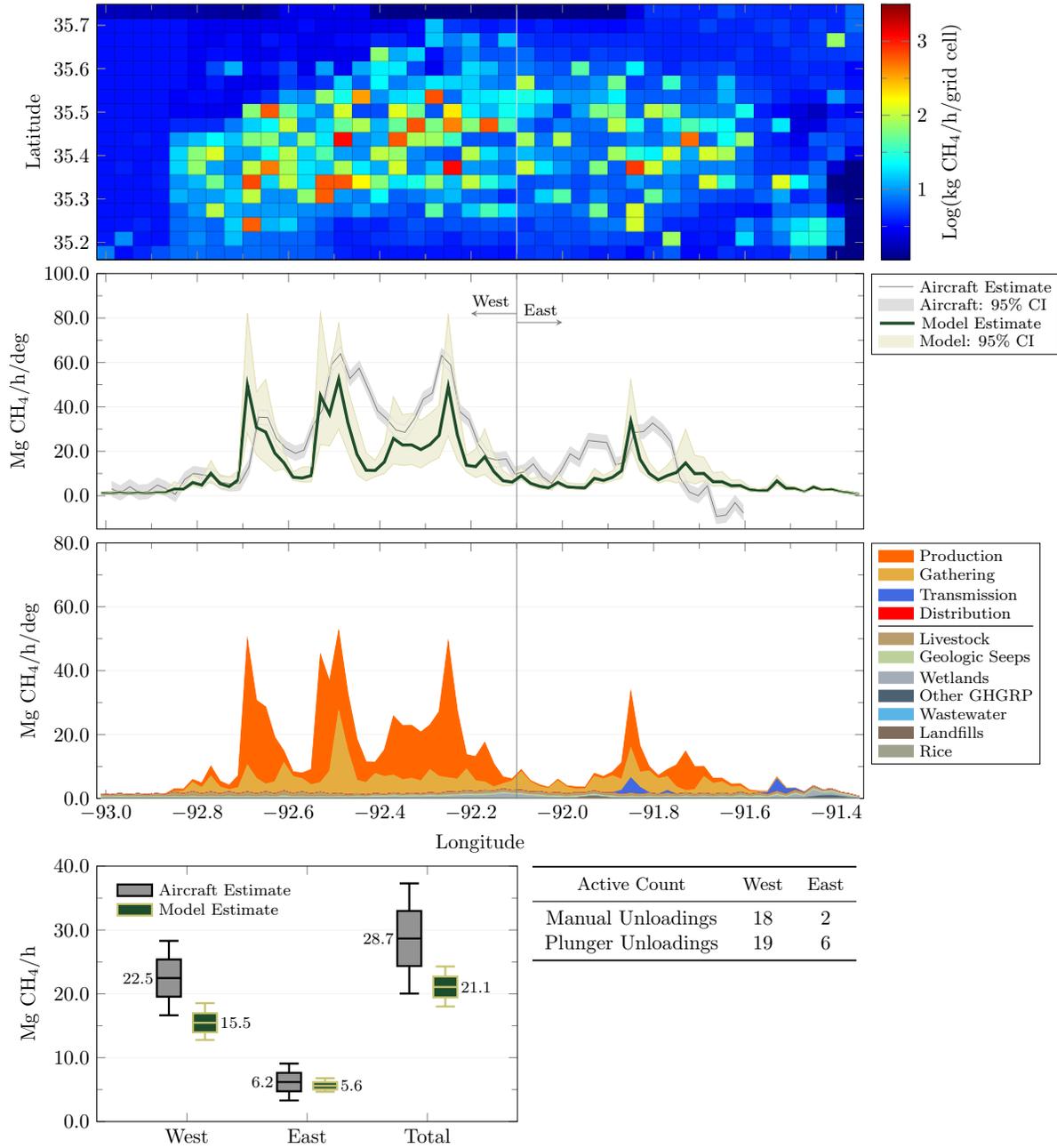
October 1, 2015 08:00-09:00 CDT



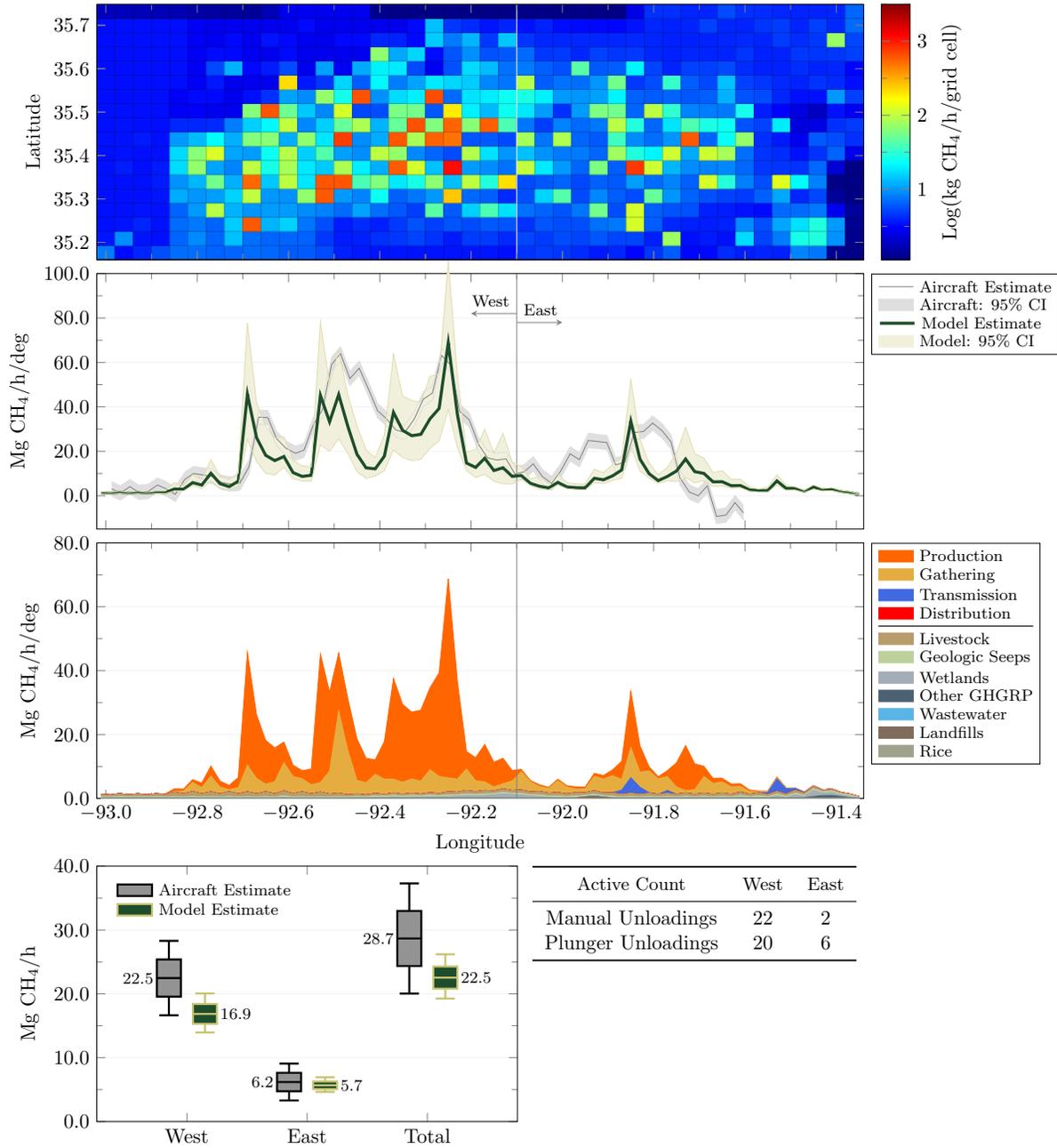
October 1, 2015 09:00-10:00 CDT



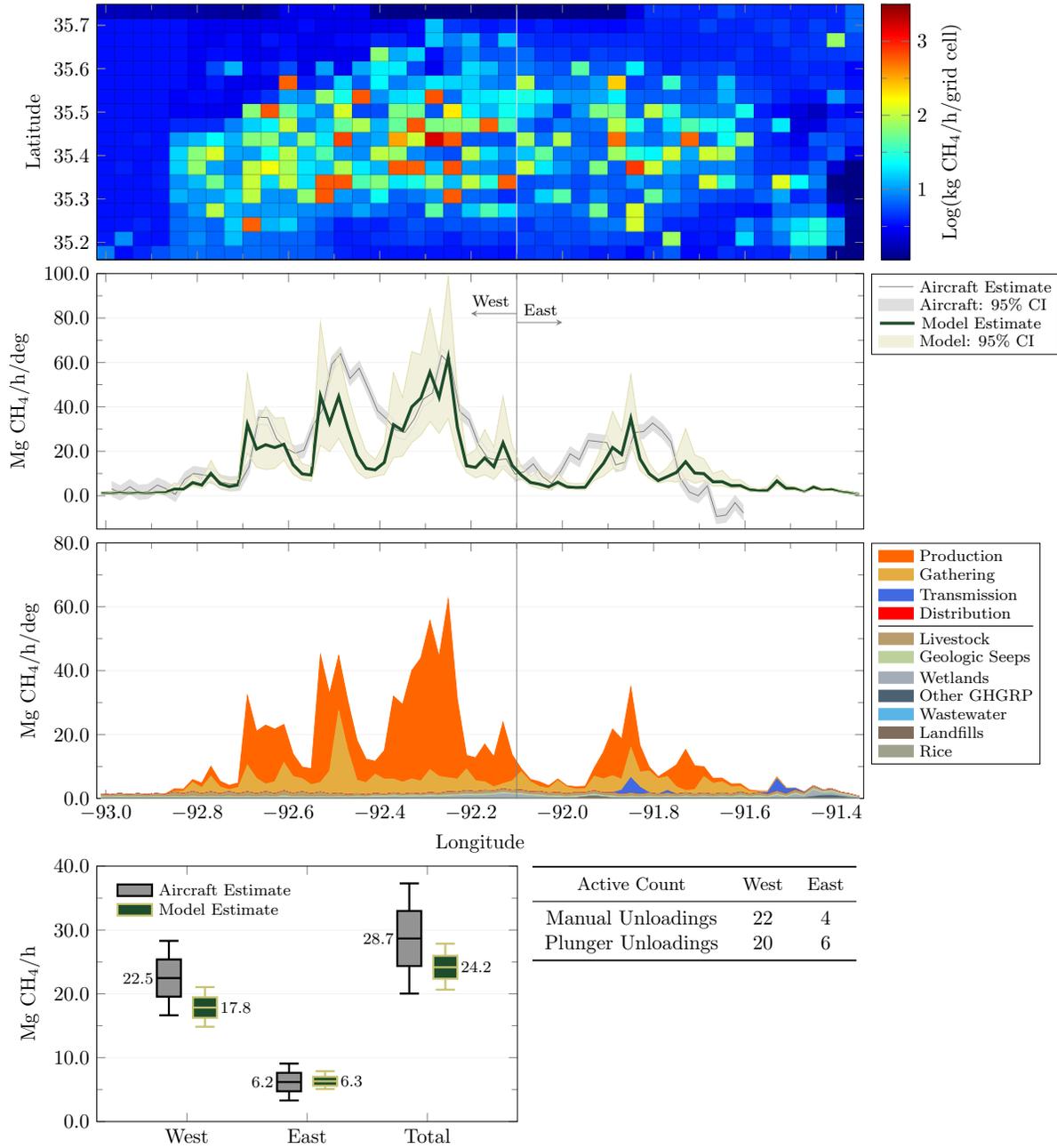
October 1, 2015 10:00-11:00 CDT



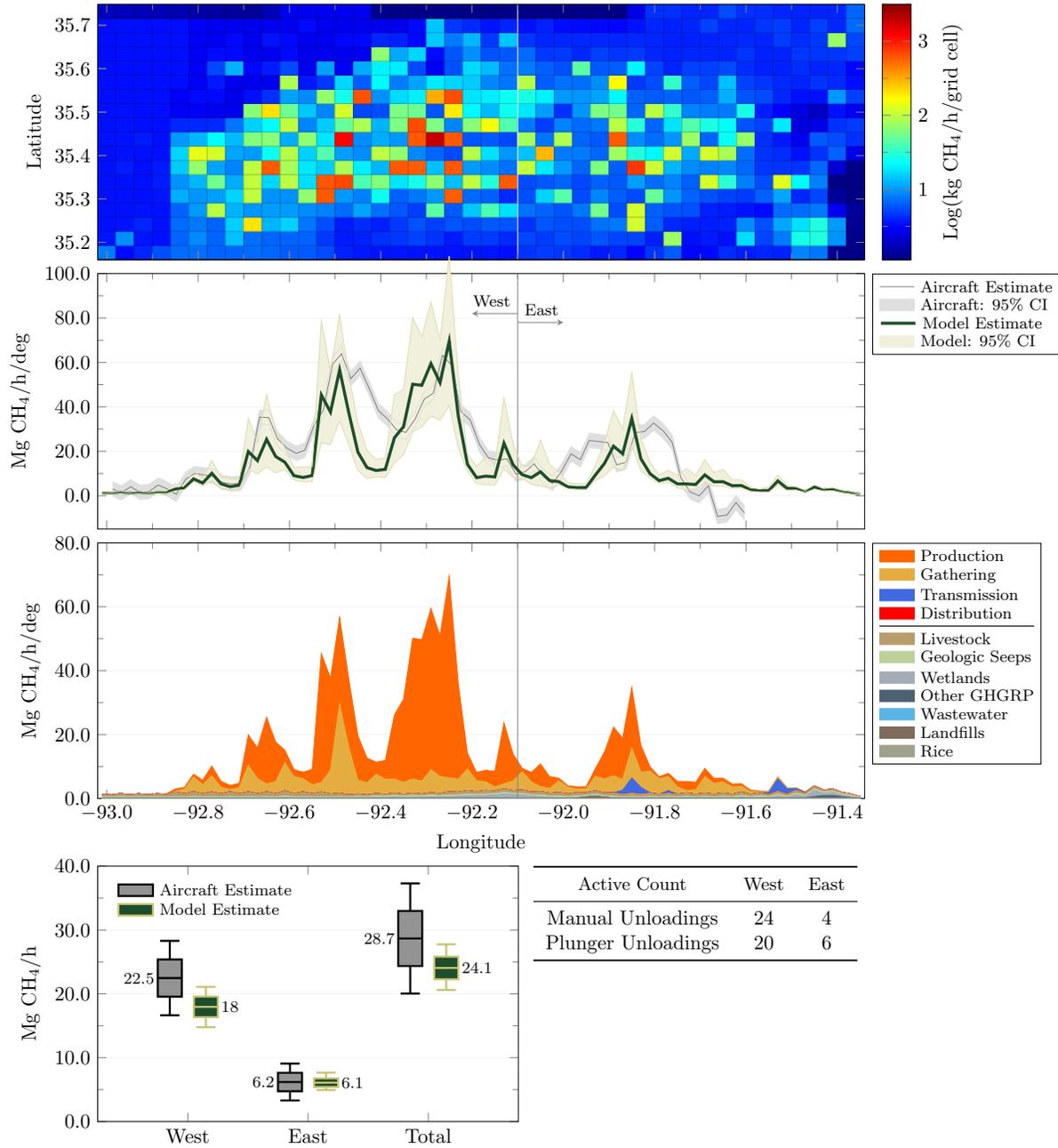
October 1, 2015 11:00-12:00 CDT



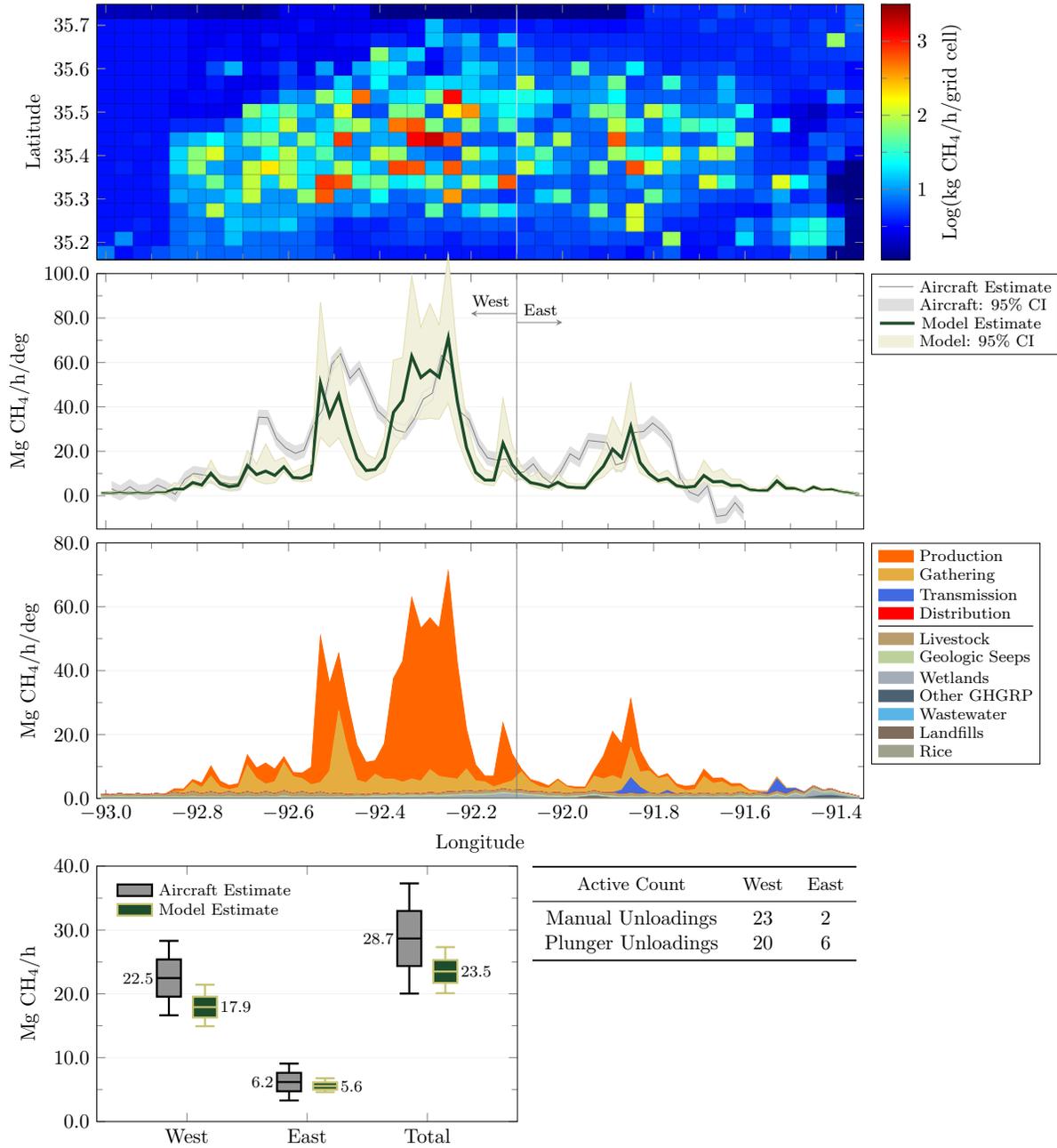
October 1, 2015 12:00-13:00 CDT



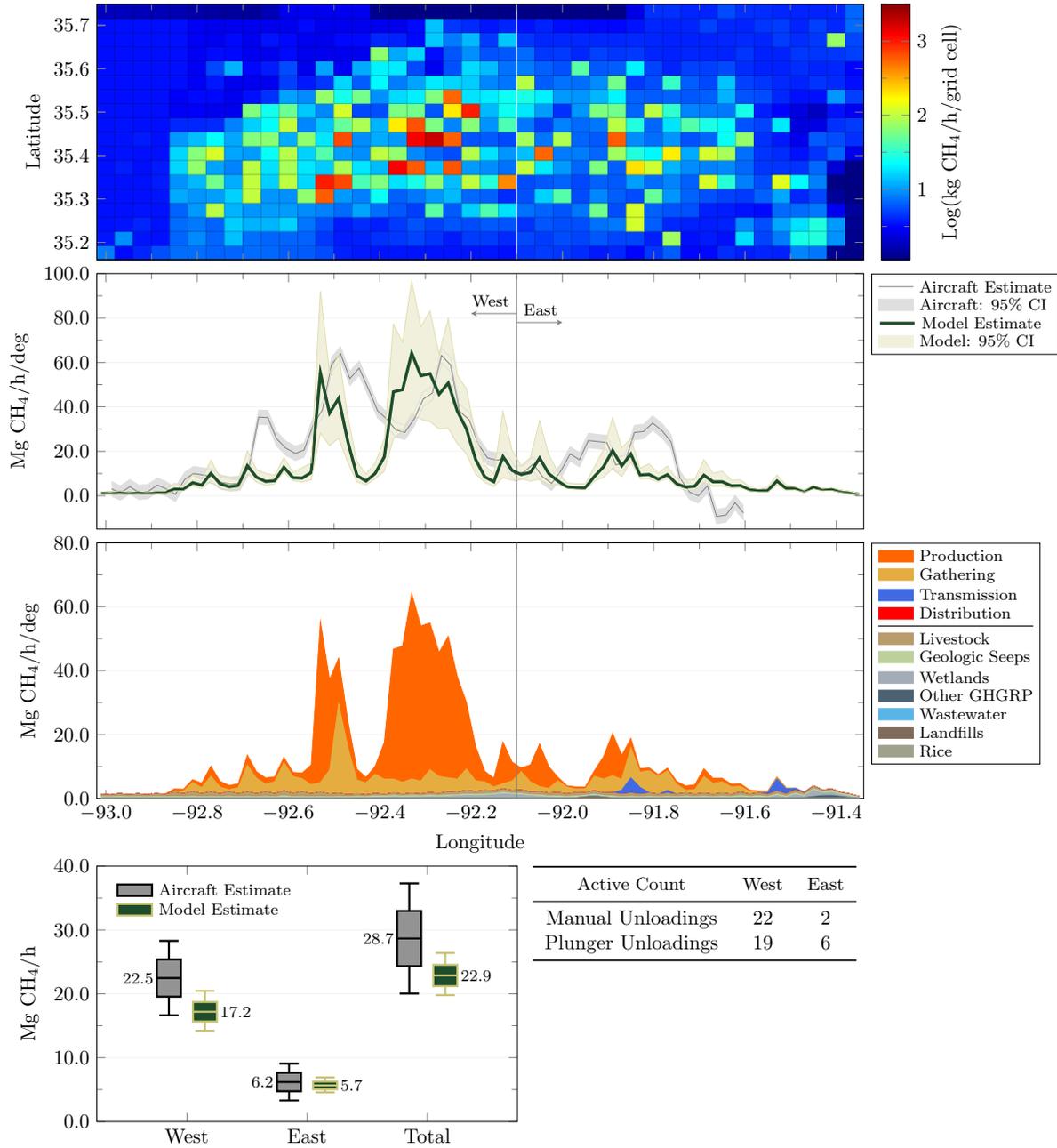
October 1, 2015 13:00-14:00 CDT



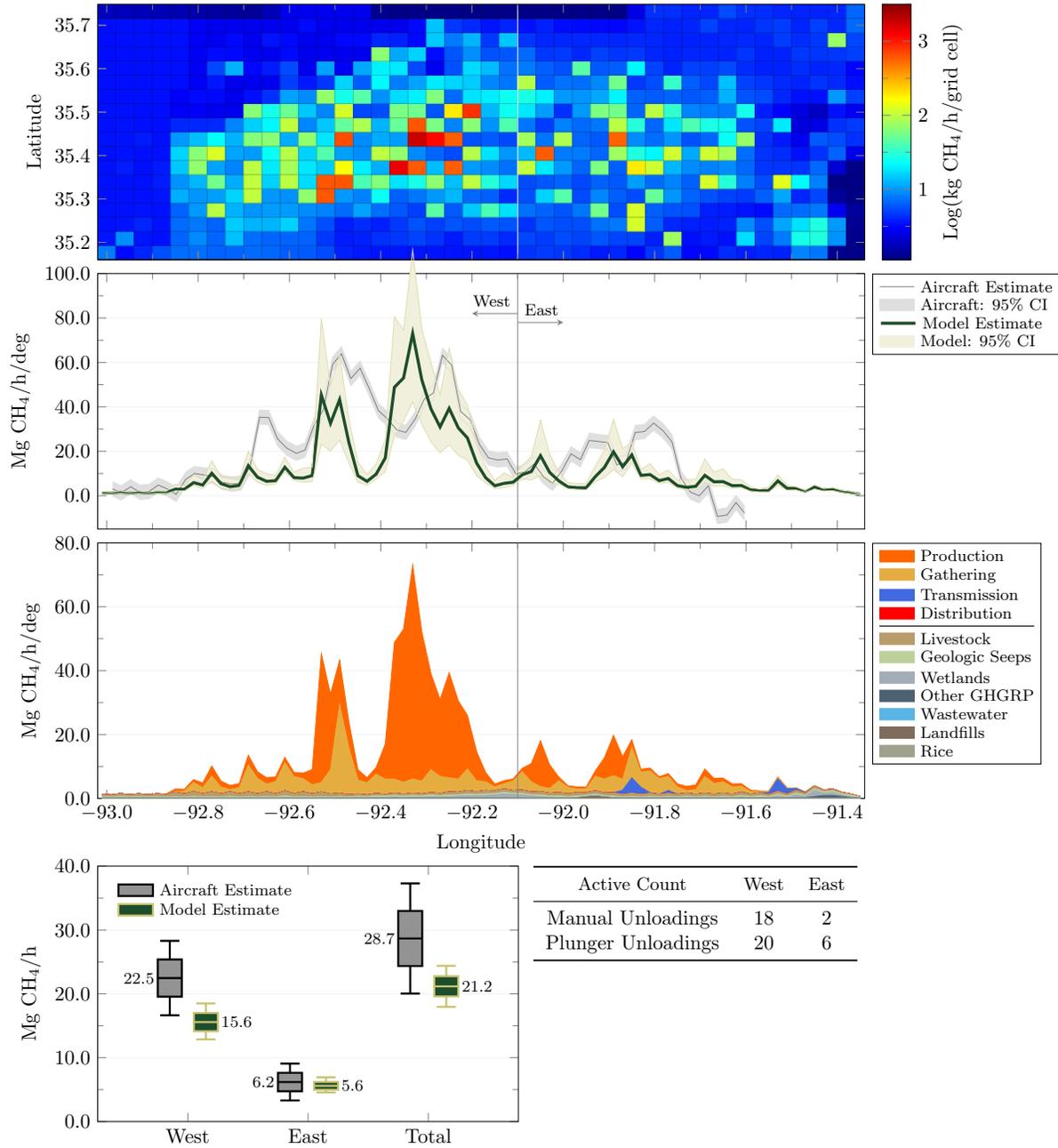
October 1, 2015 14:00-15:00 CDT



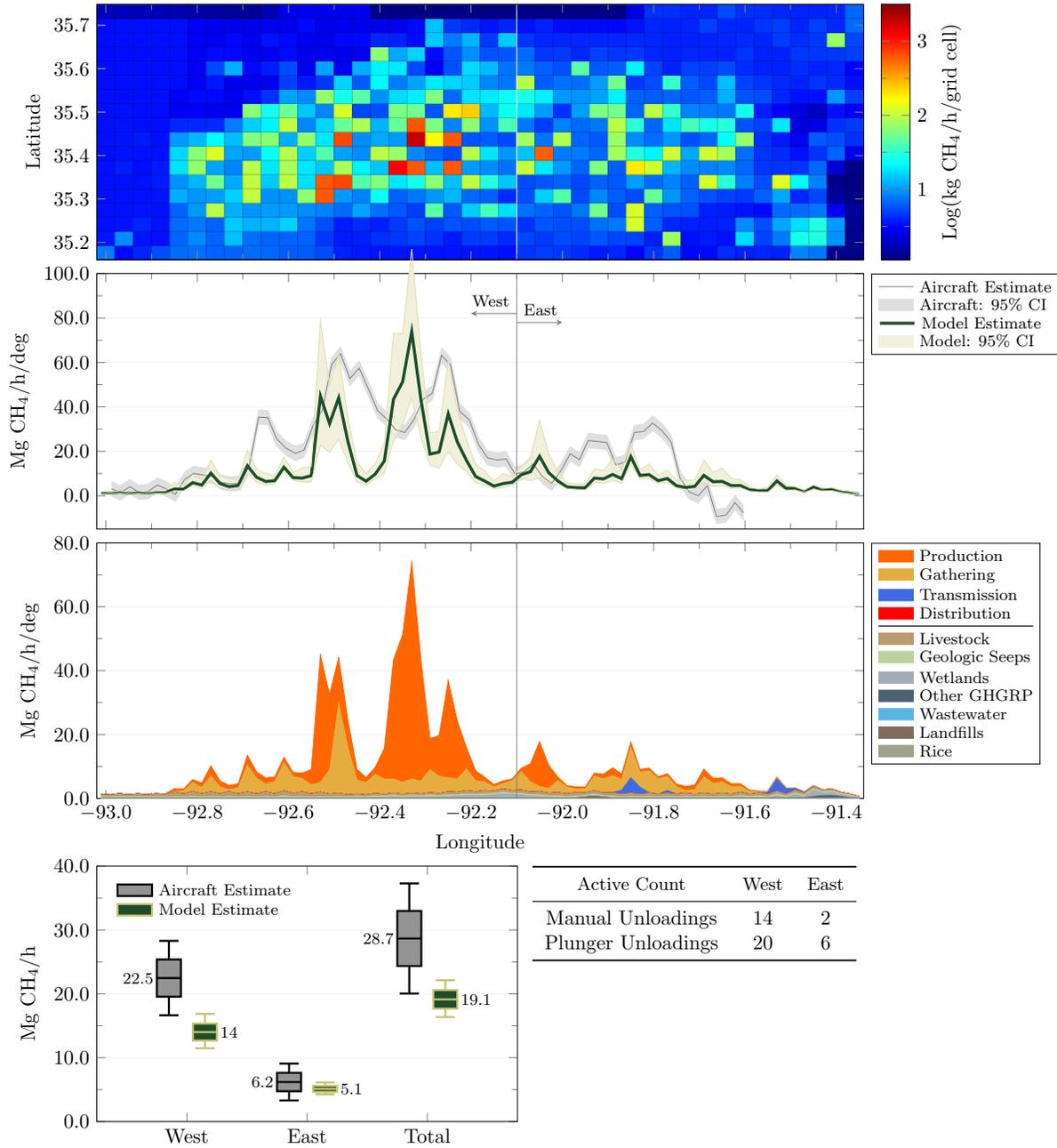
October 1, 2015 15:00-16:00 CDT



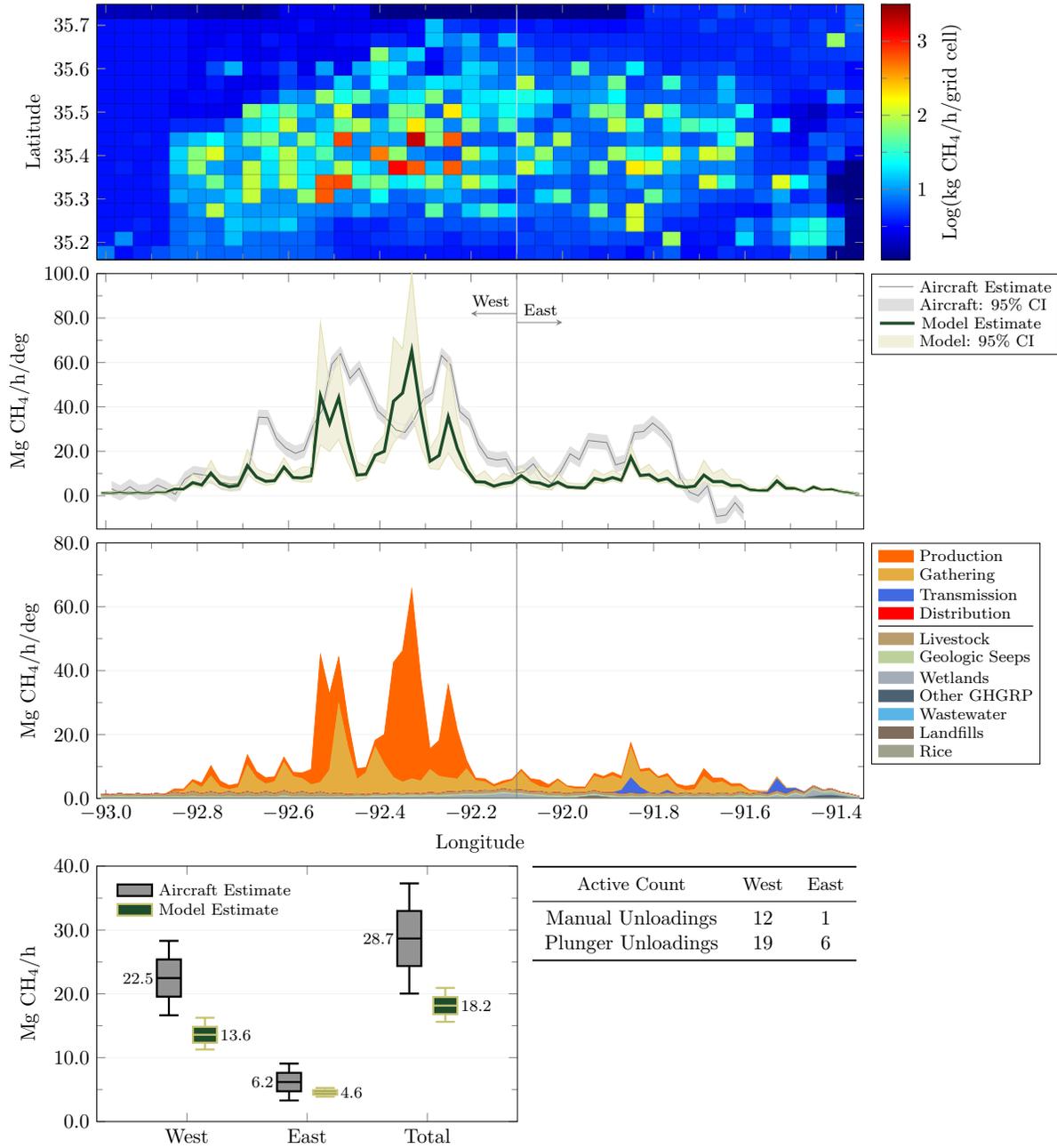
October 1, 2015 16:00-17:00 CDT



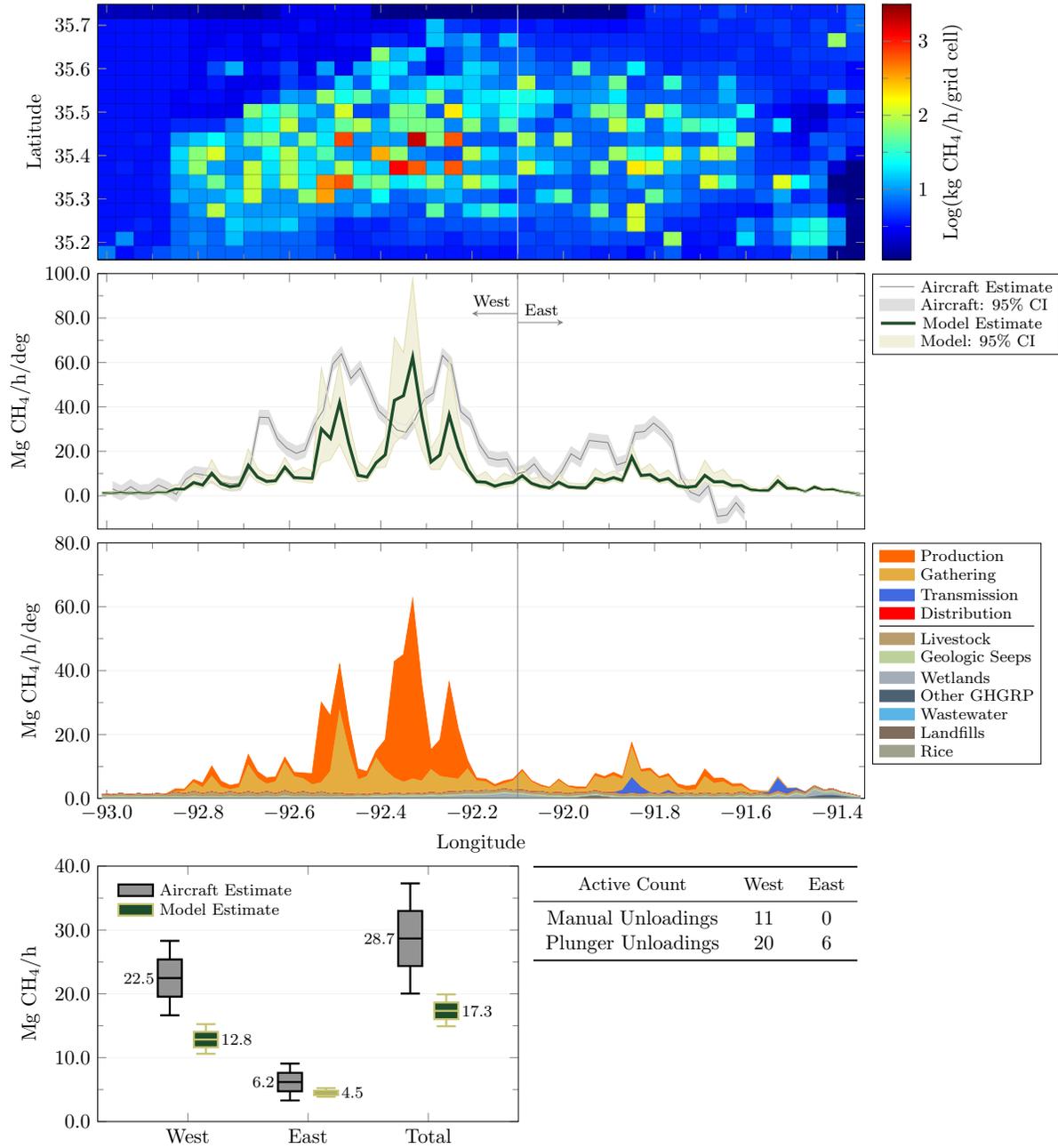
October 1, 2015 17:00-18:00 CDT



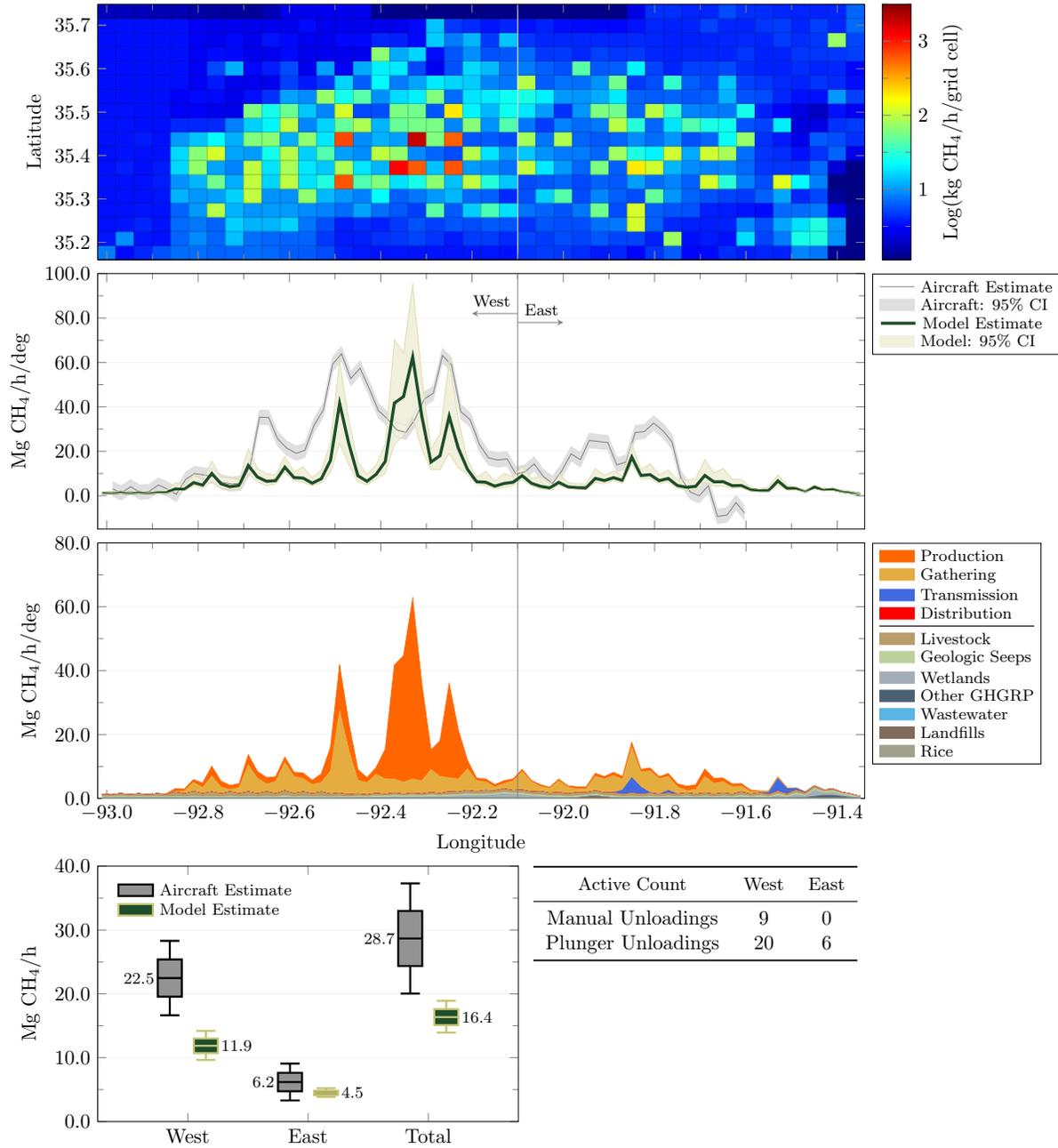
October 1, 2015 18:00-19:00 CDT



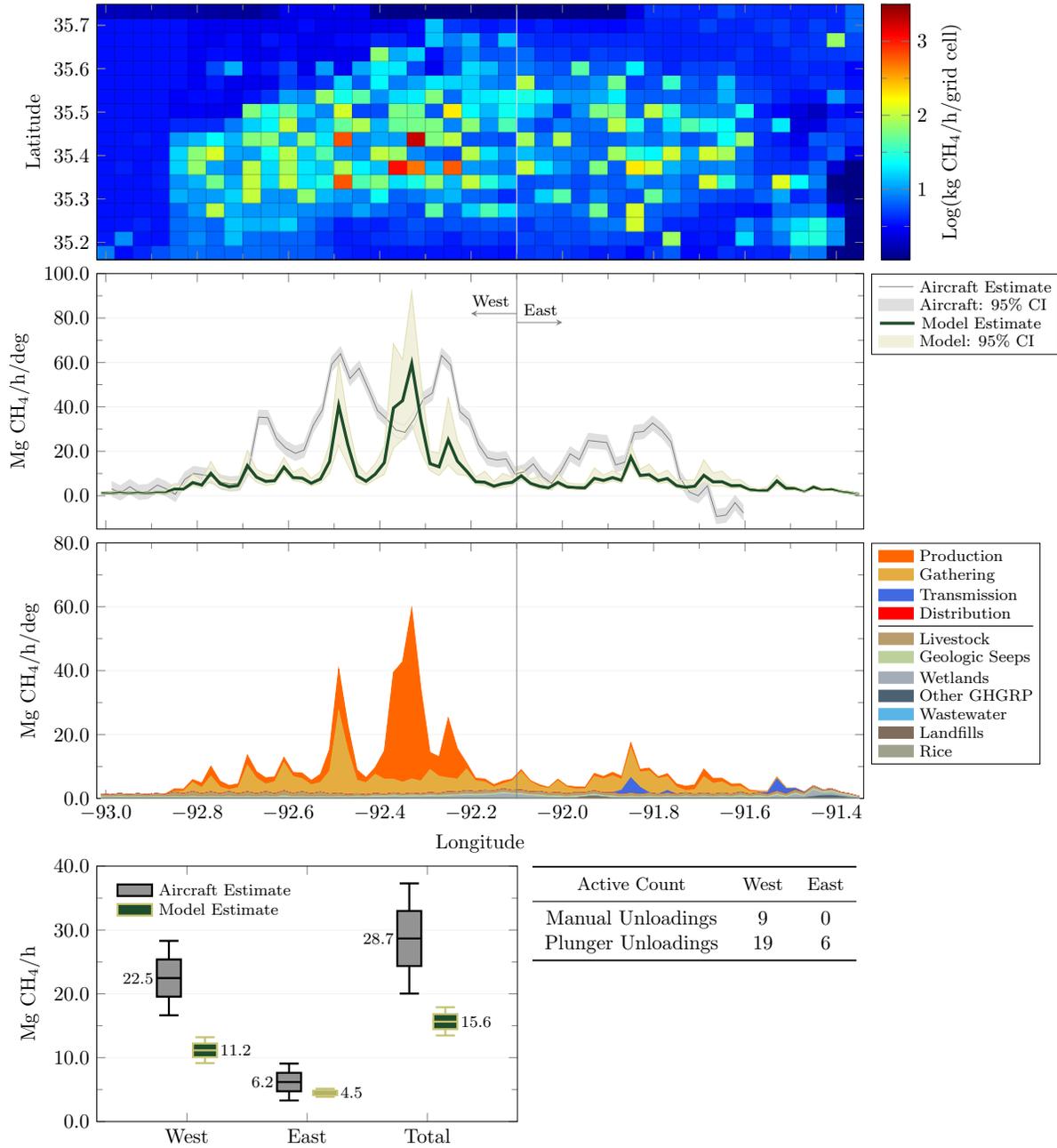
October 1, 2015 19:00-20:00 CDT



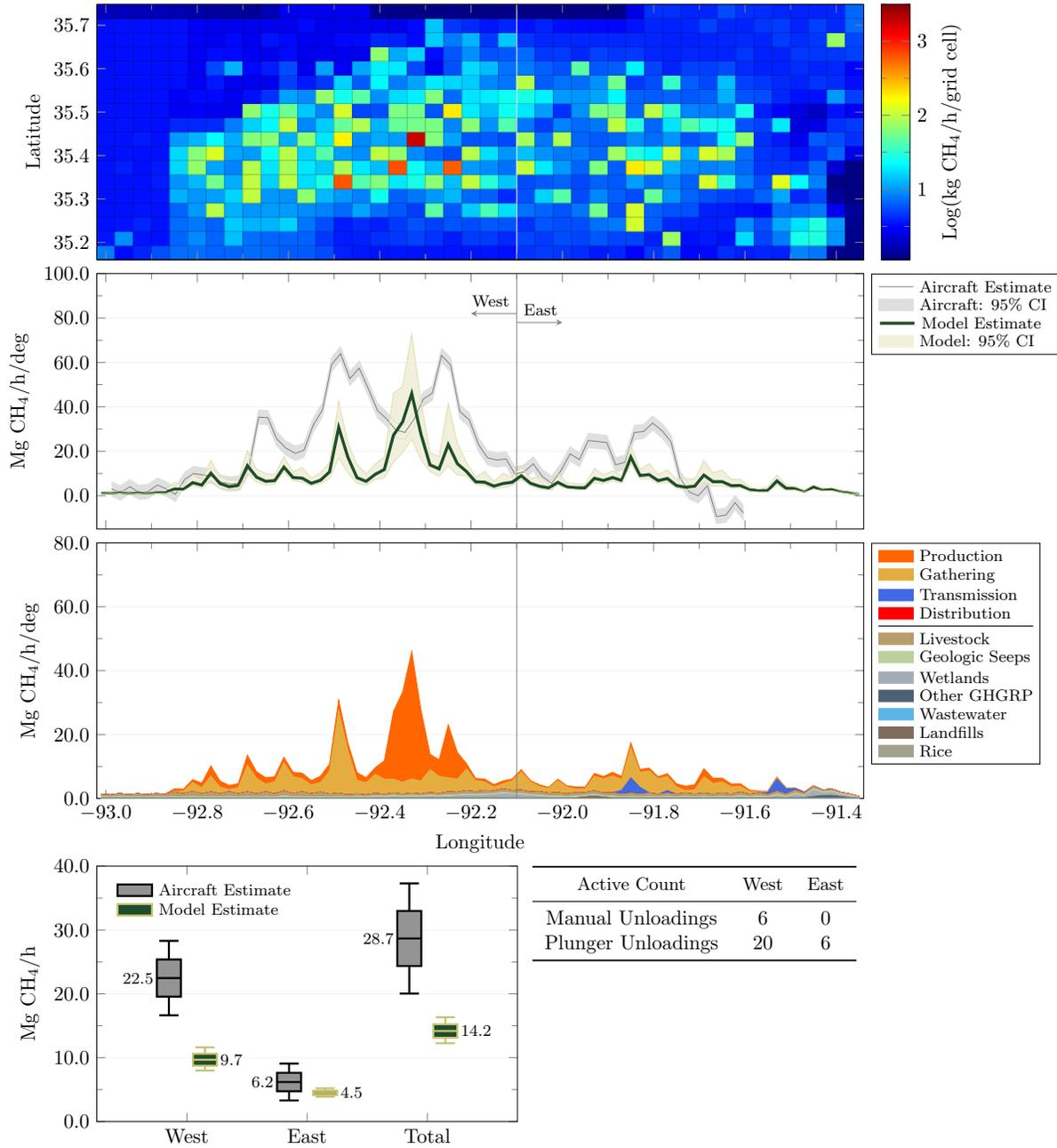
October 1, 2015 20:00-21:00 CDT



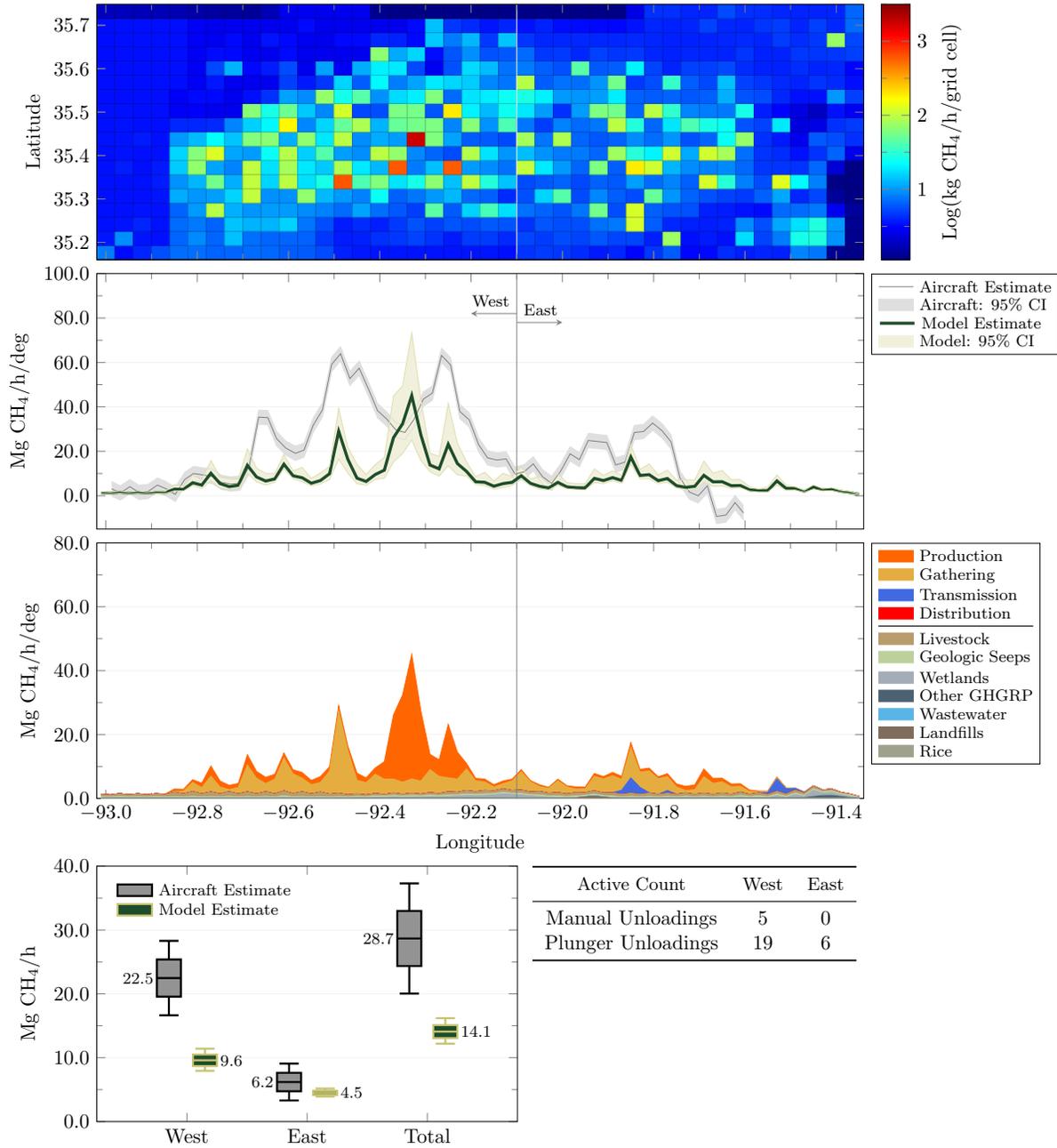
October 1, 2015 21:00-22:00 CDT



October 1, 2015 22:00-23:00 CDT

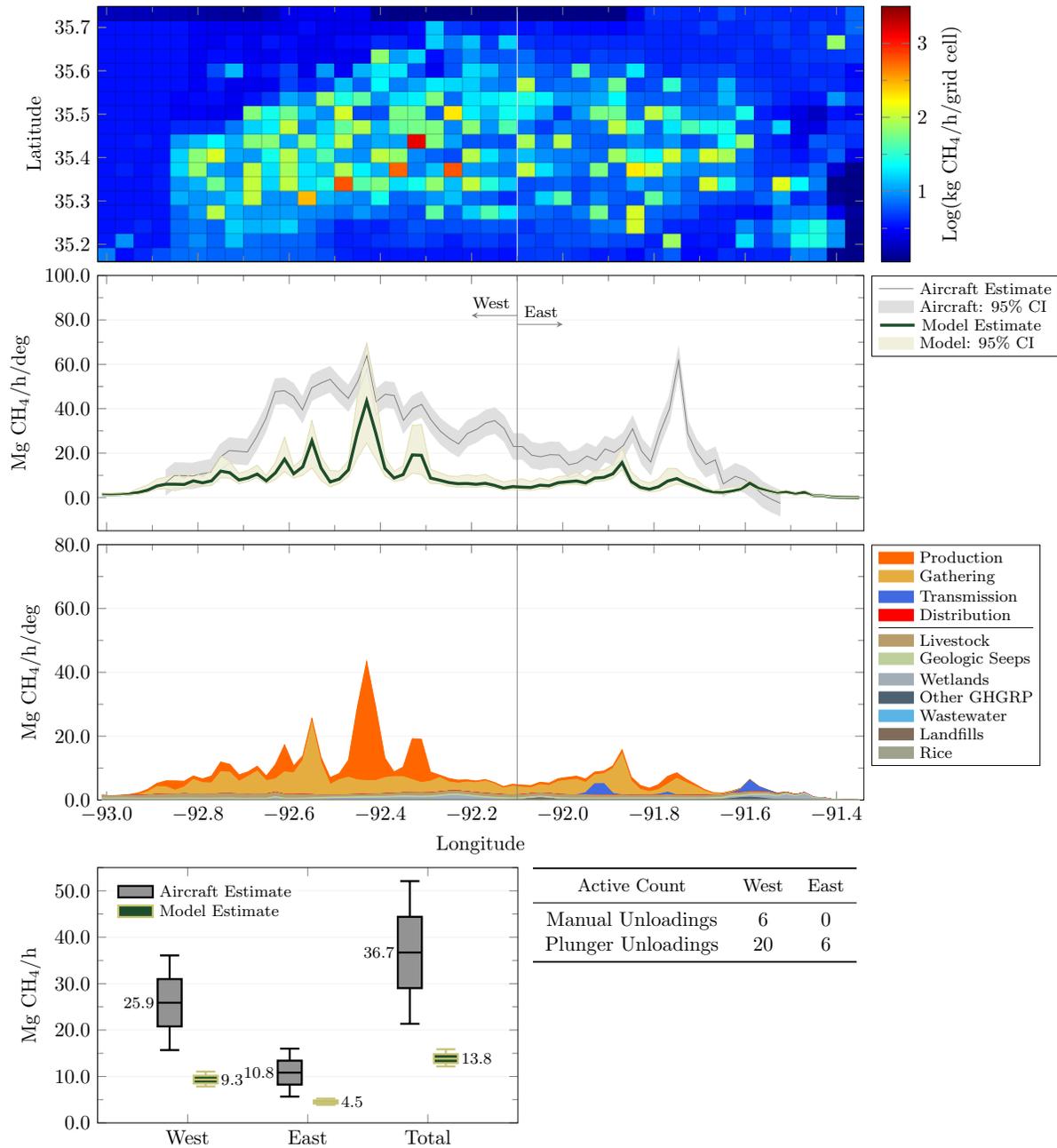


October 1, 2015 23:00-00:00 CDT

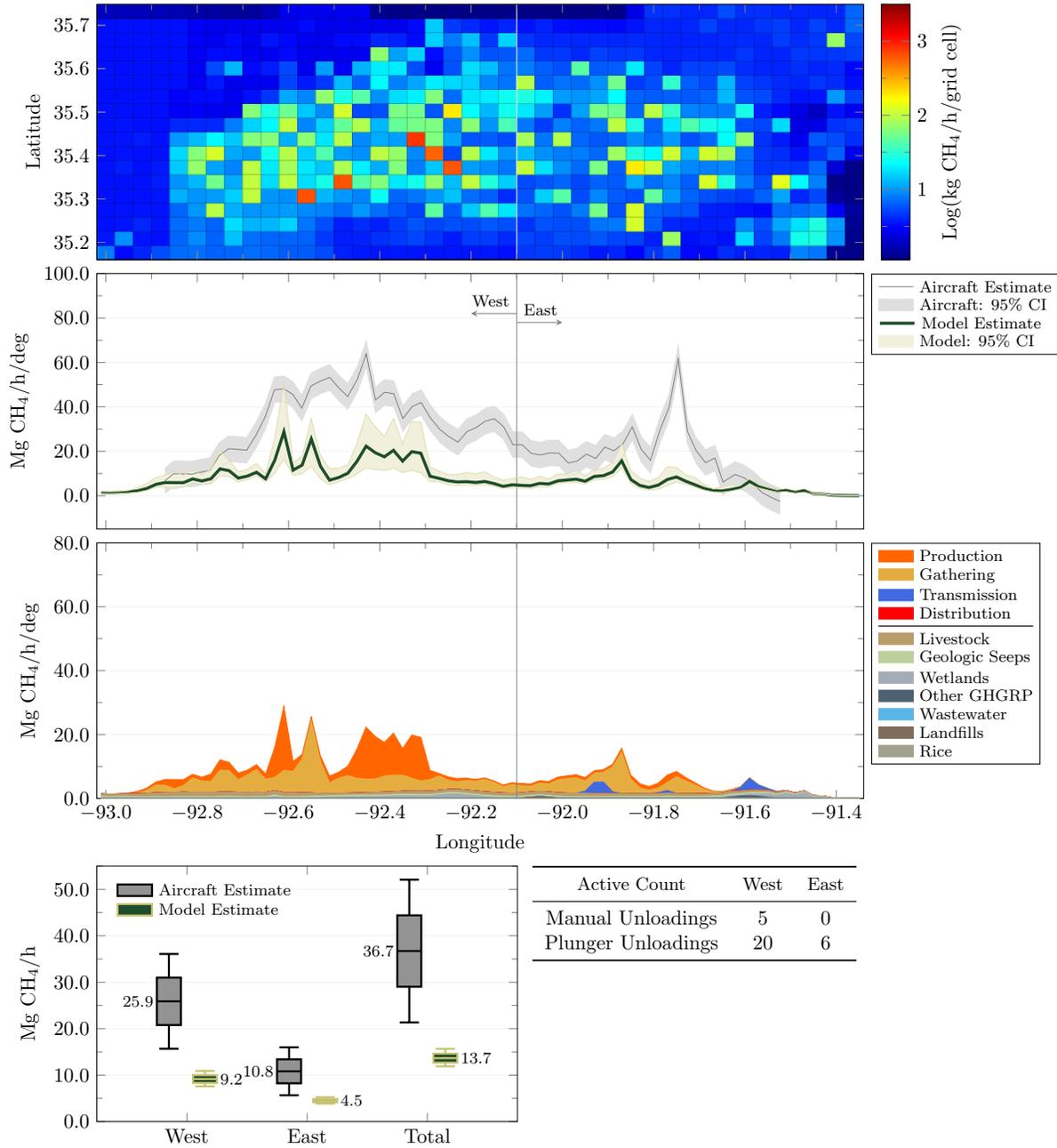


C.3.2 HOURLY GLAE RESULTS: OCTOBER 2ND, 2015

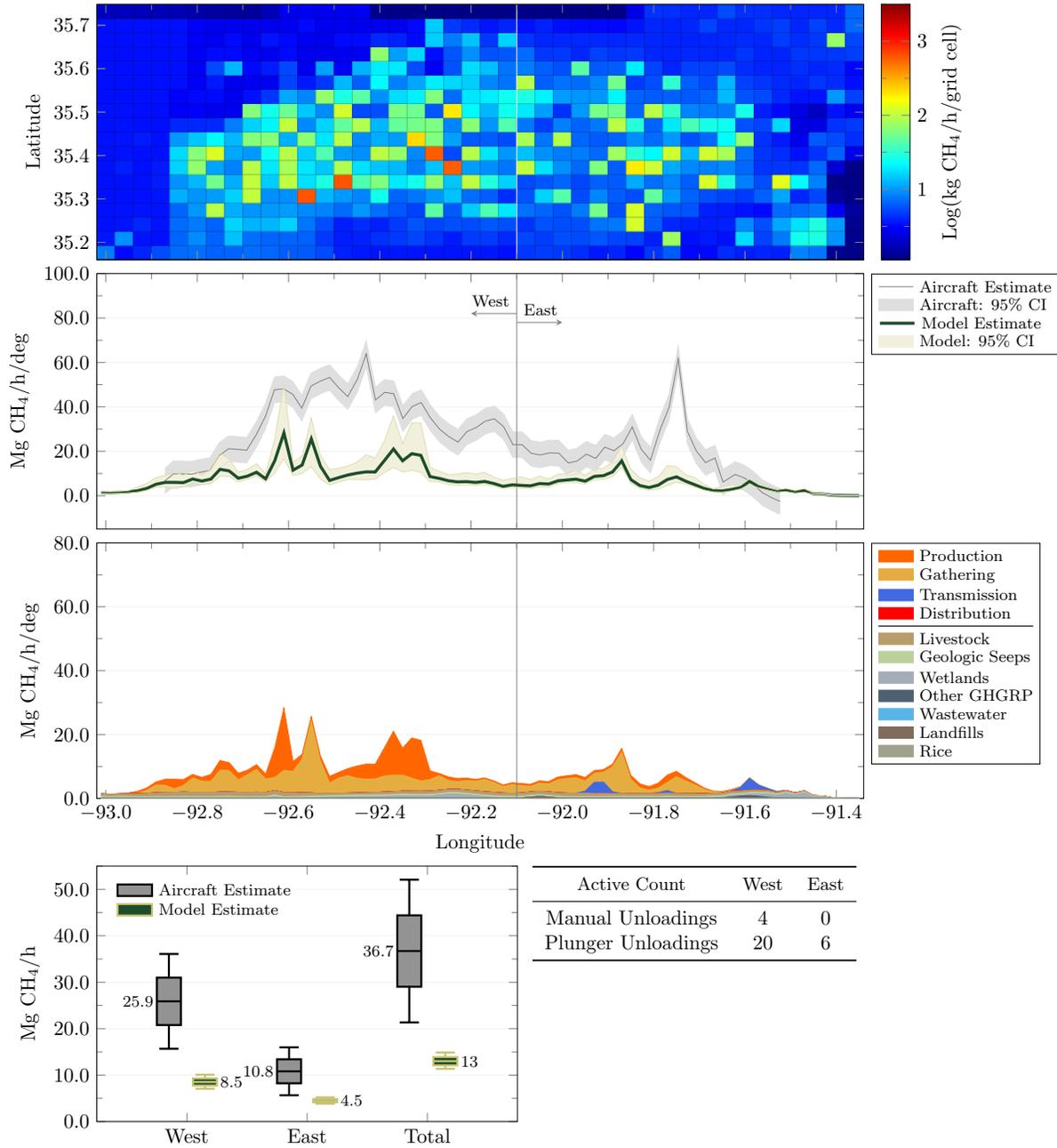
October 2, 2015 00:00-01:00 CDT



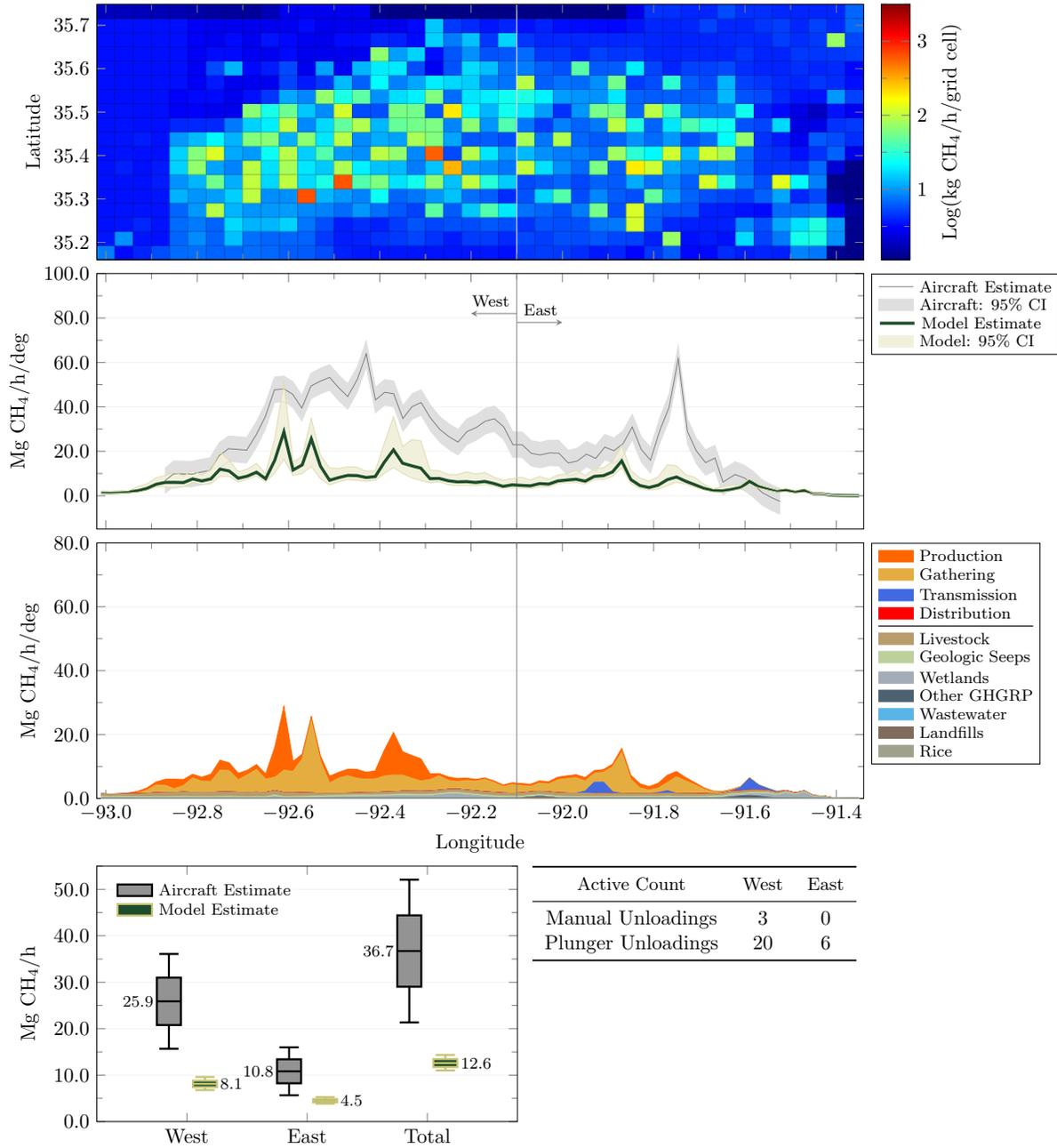
October 2, 2015 01:00-02:00 CDT



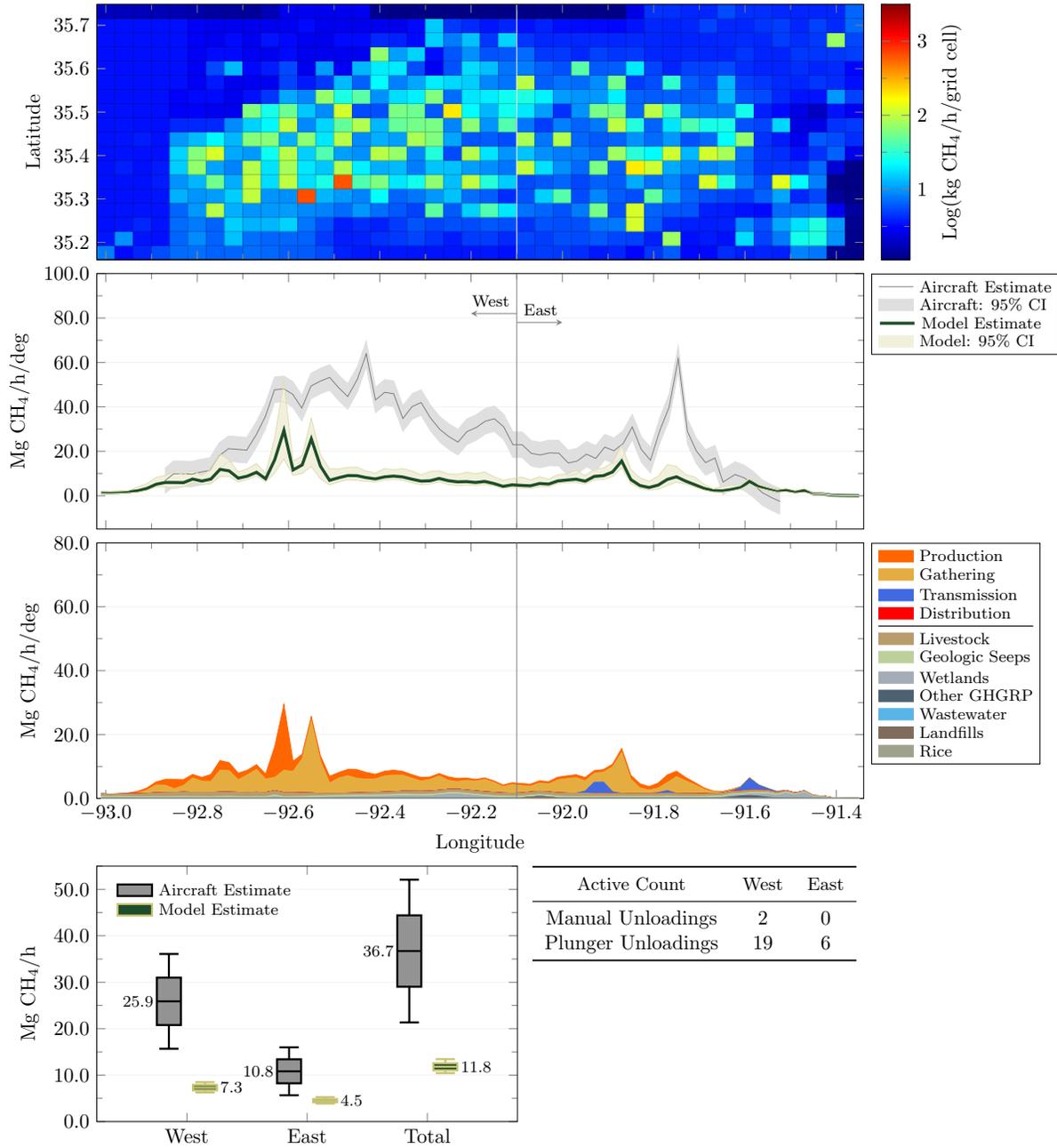
October 2, 2015 02:00-03:00 CDT



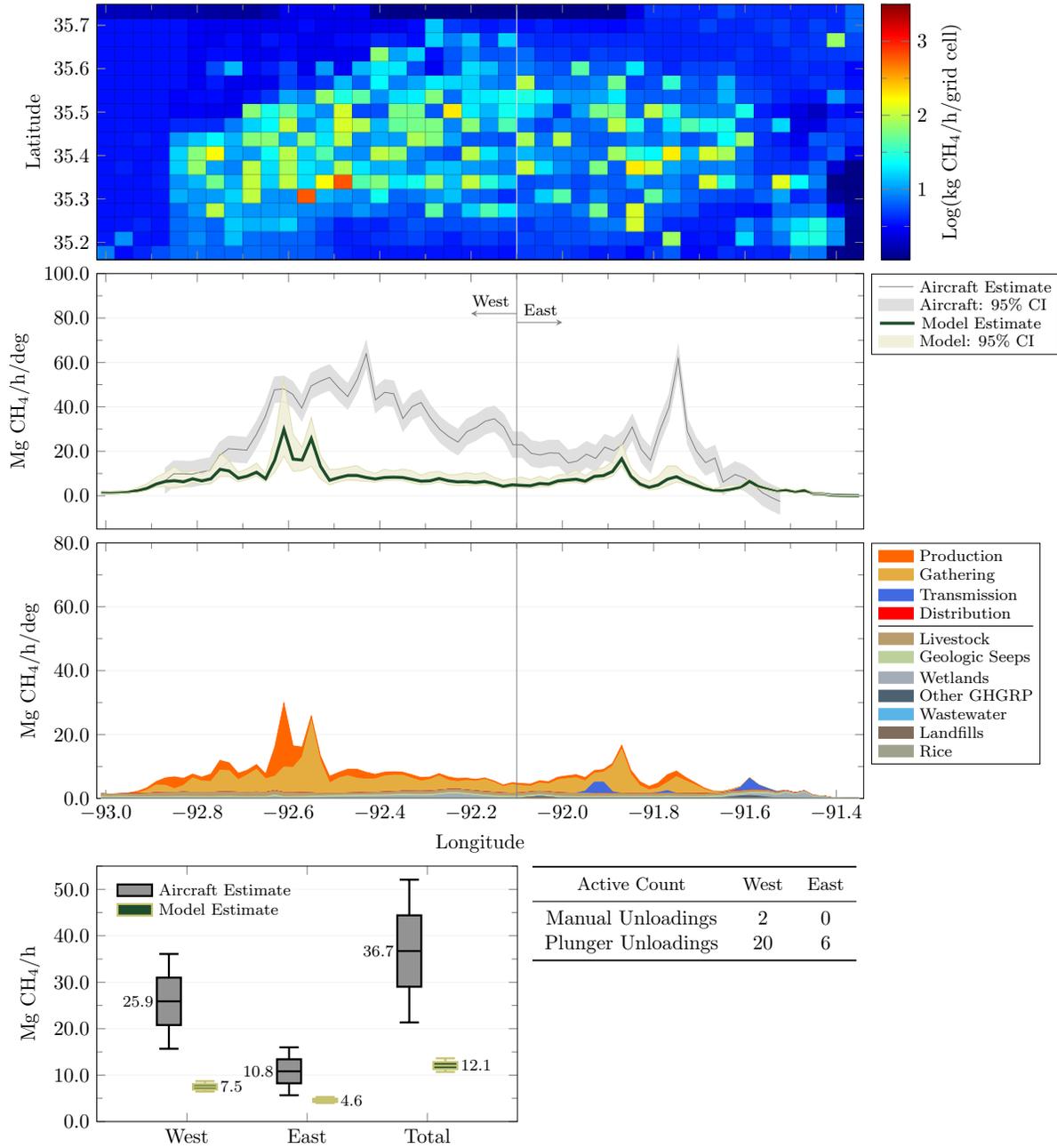
October 2, 2015 03:00-04:00 CDT



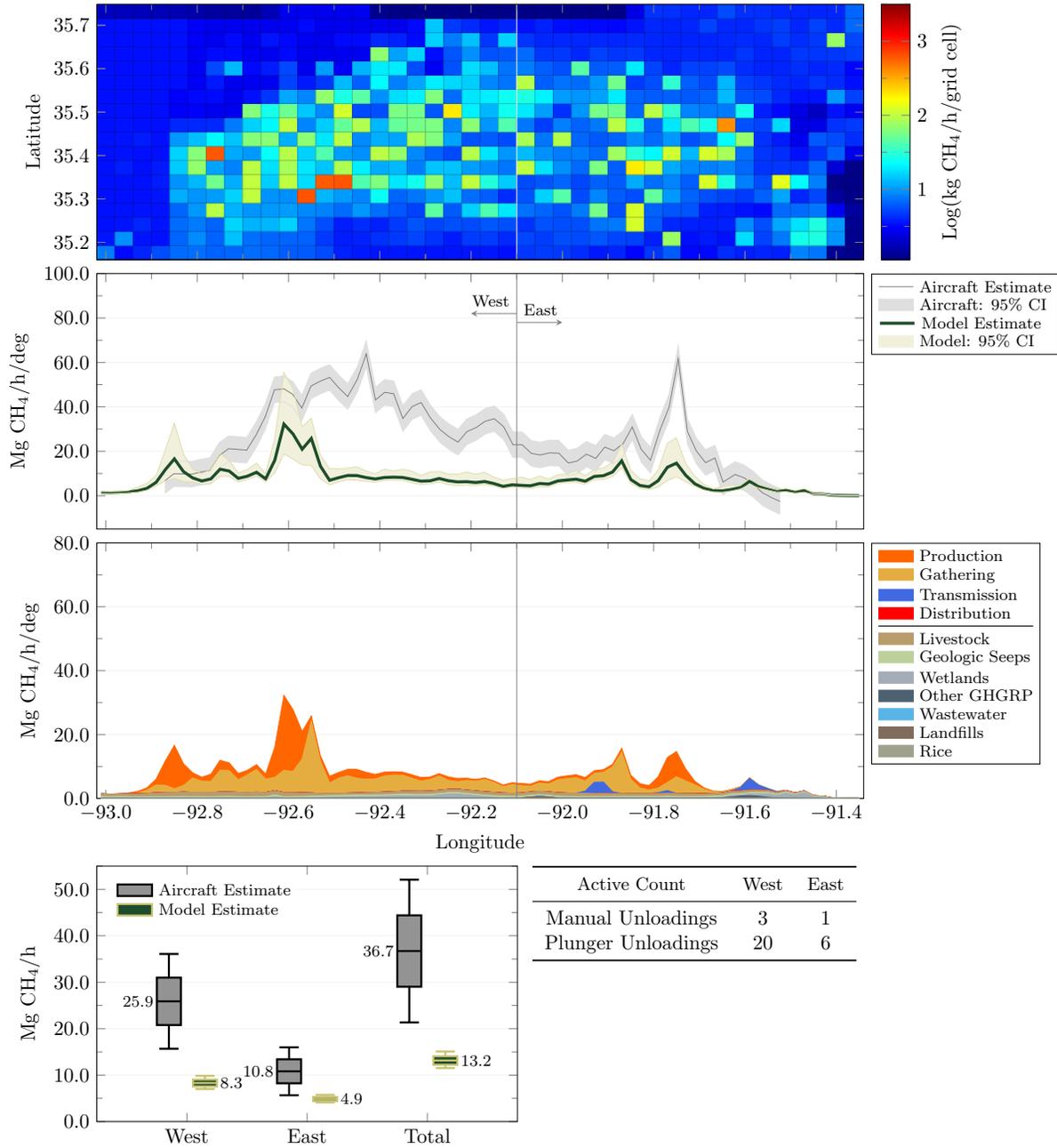
October 2, 2015 04:00-05:00 CDT



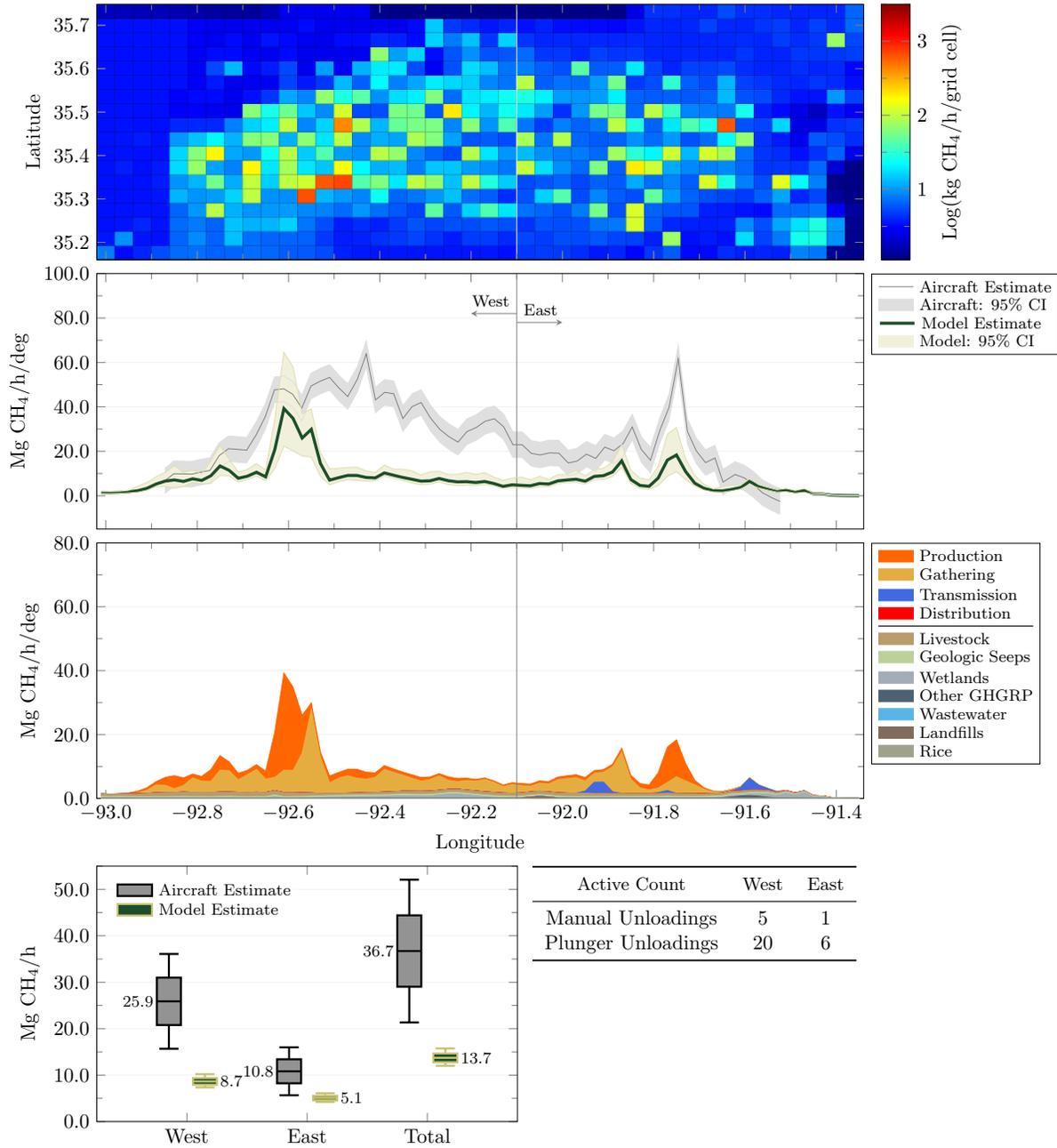
October 2, 2015 05:00-06:00 CDT



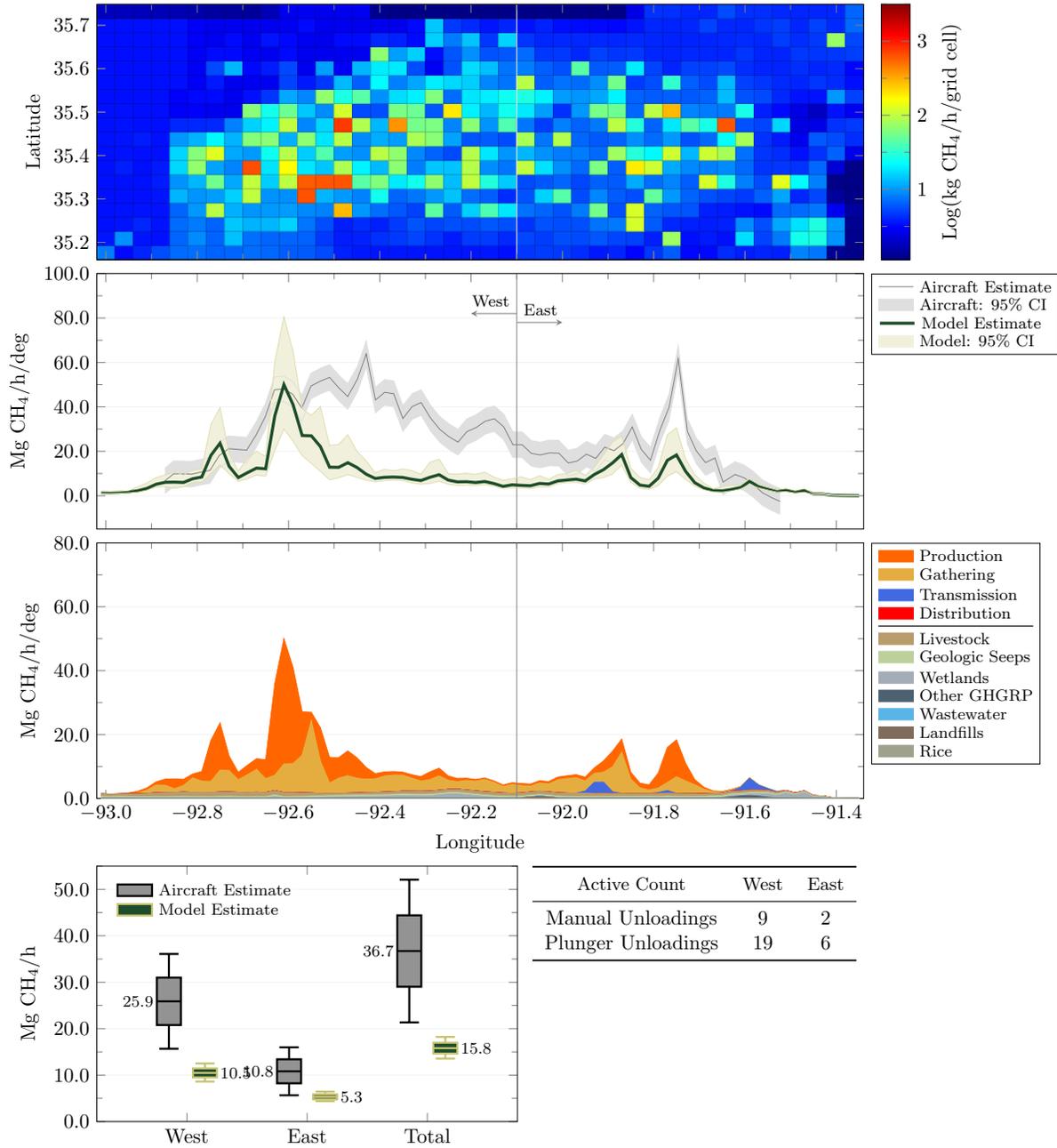
October 2, 2015 06:00-07:00 CDT



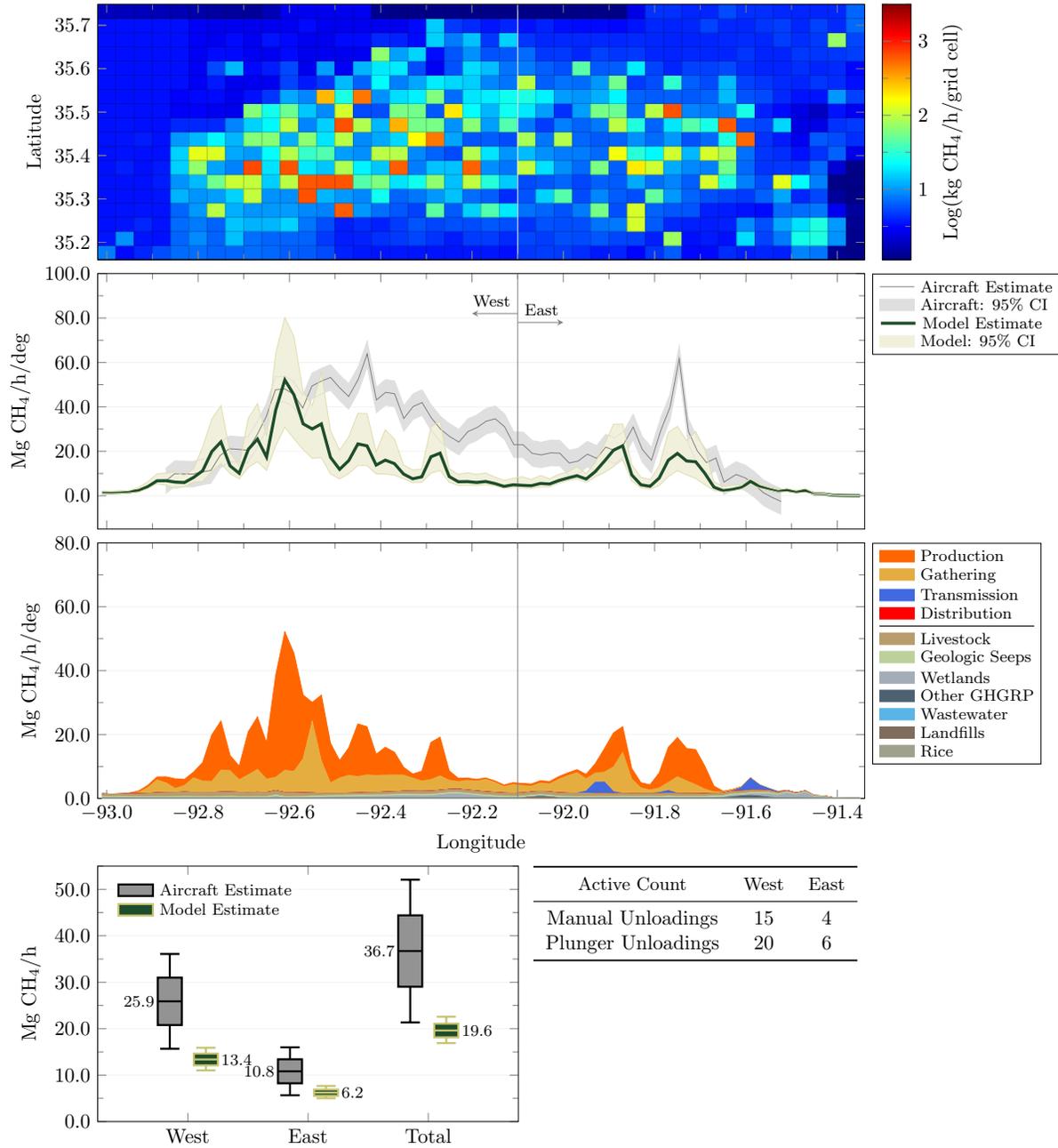
October 2, 2015 07:00-08:00 CDT



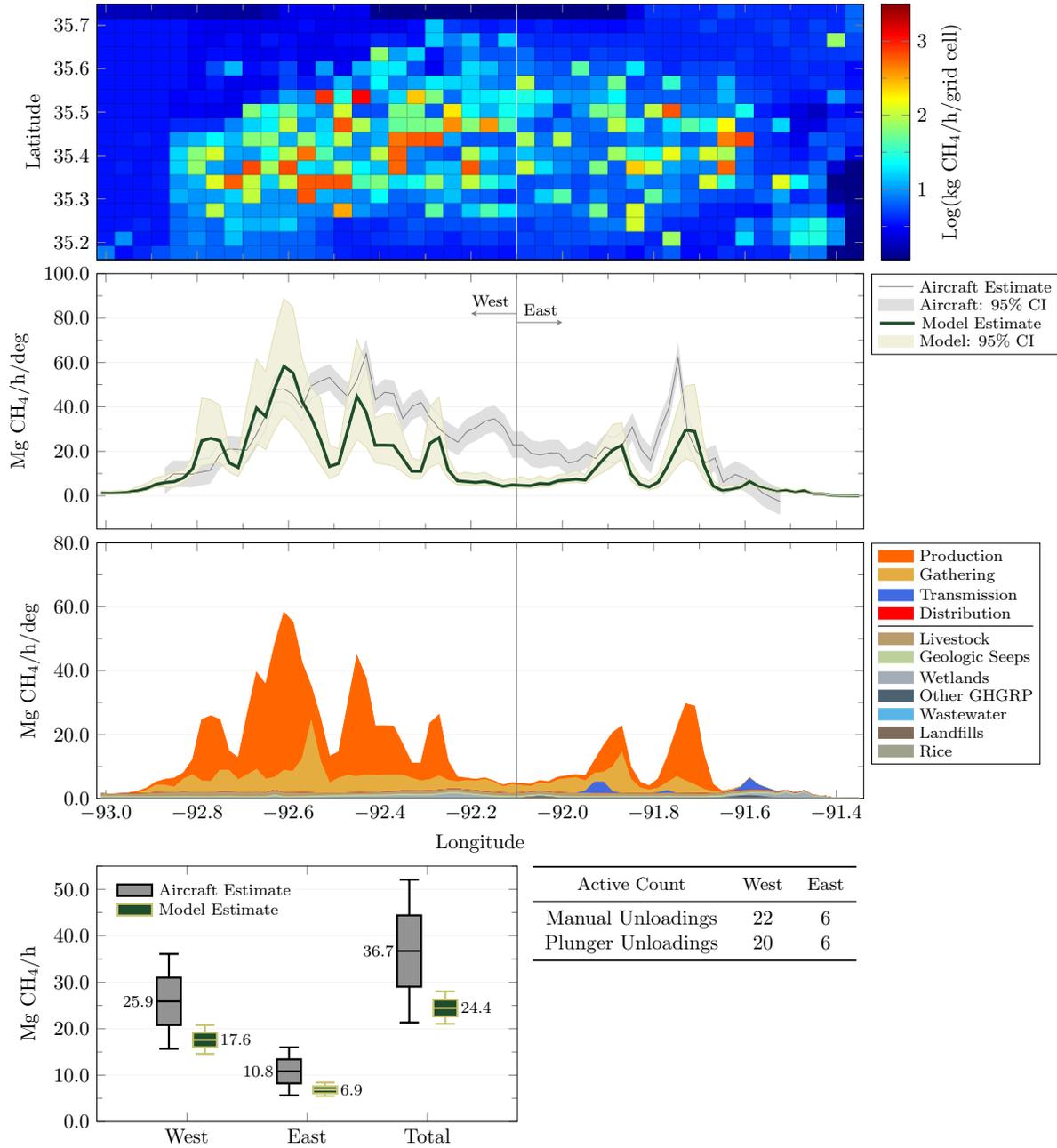
October 2, 2015 08:00-09:00 CDT



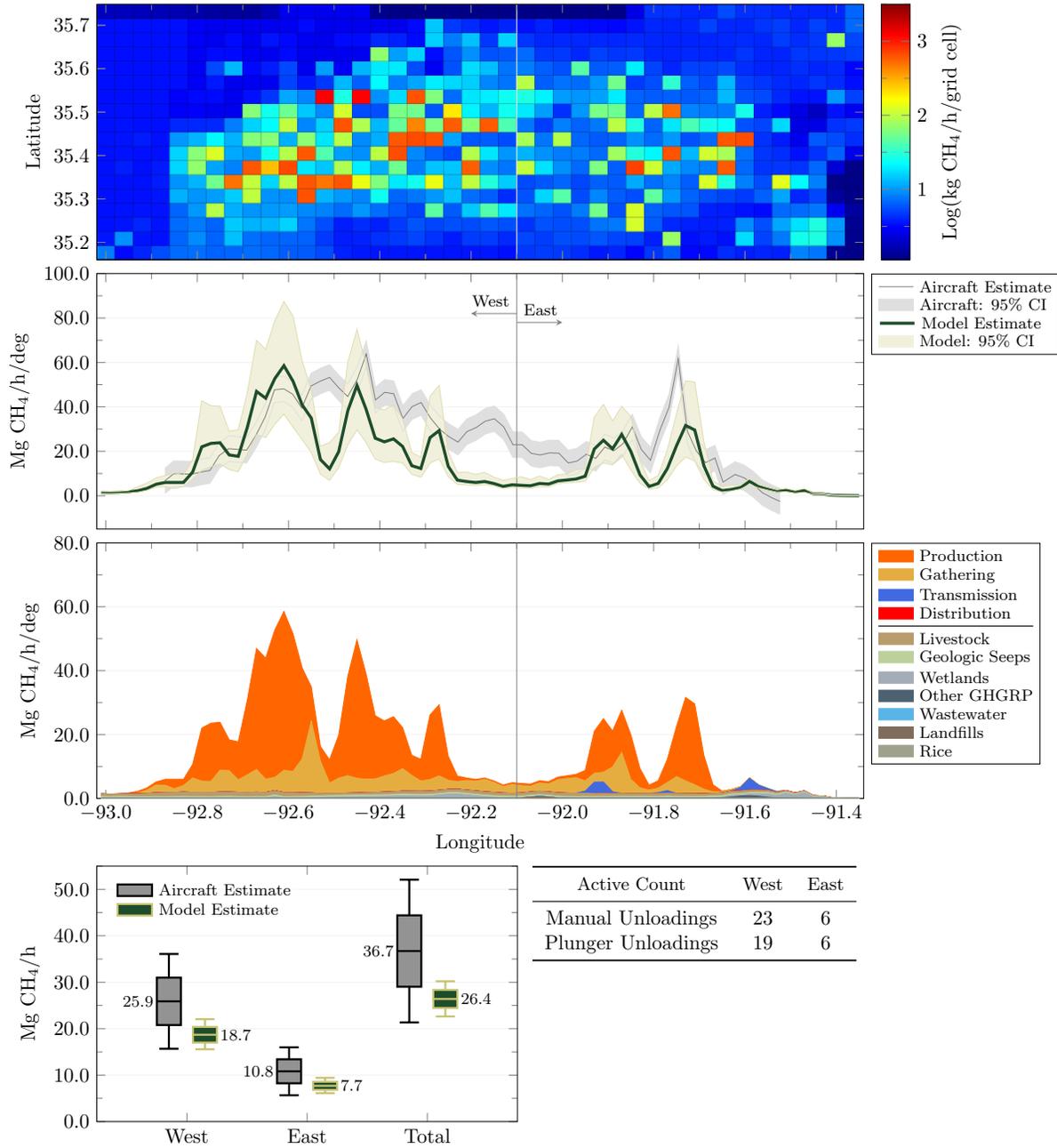
October 2, 2015 09:00-10:00 CDT



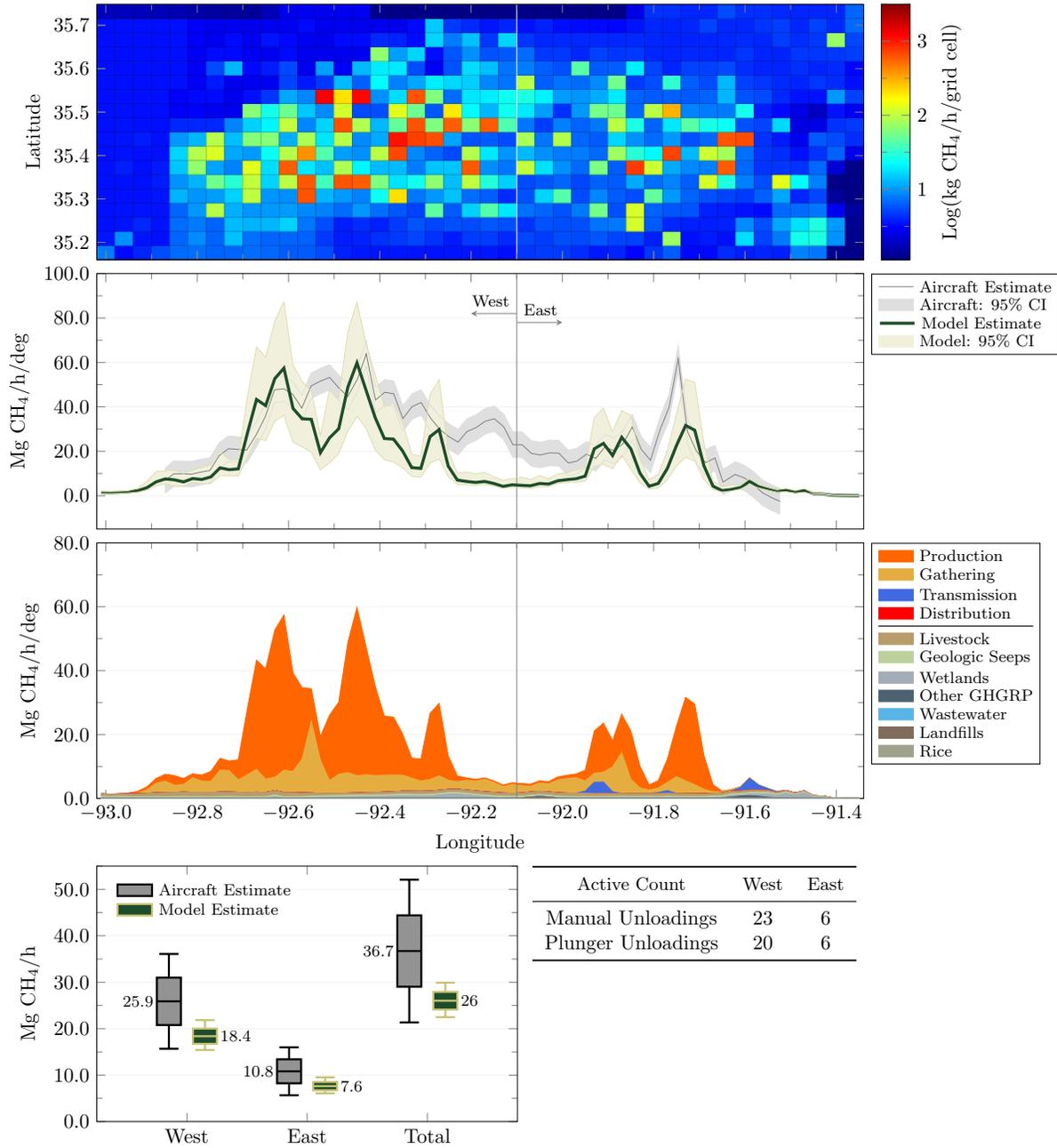
October 2, 2015 10:00-11:00 CDT



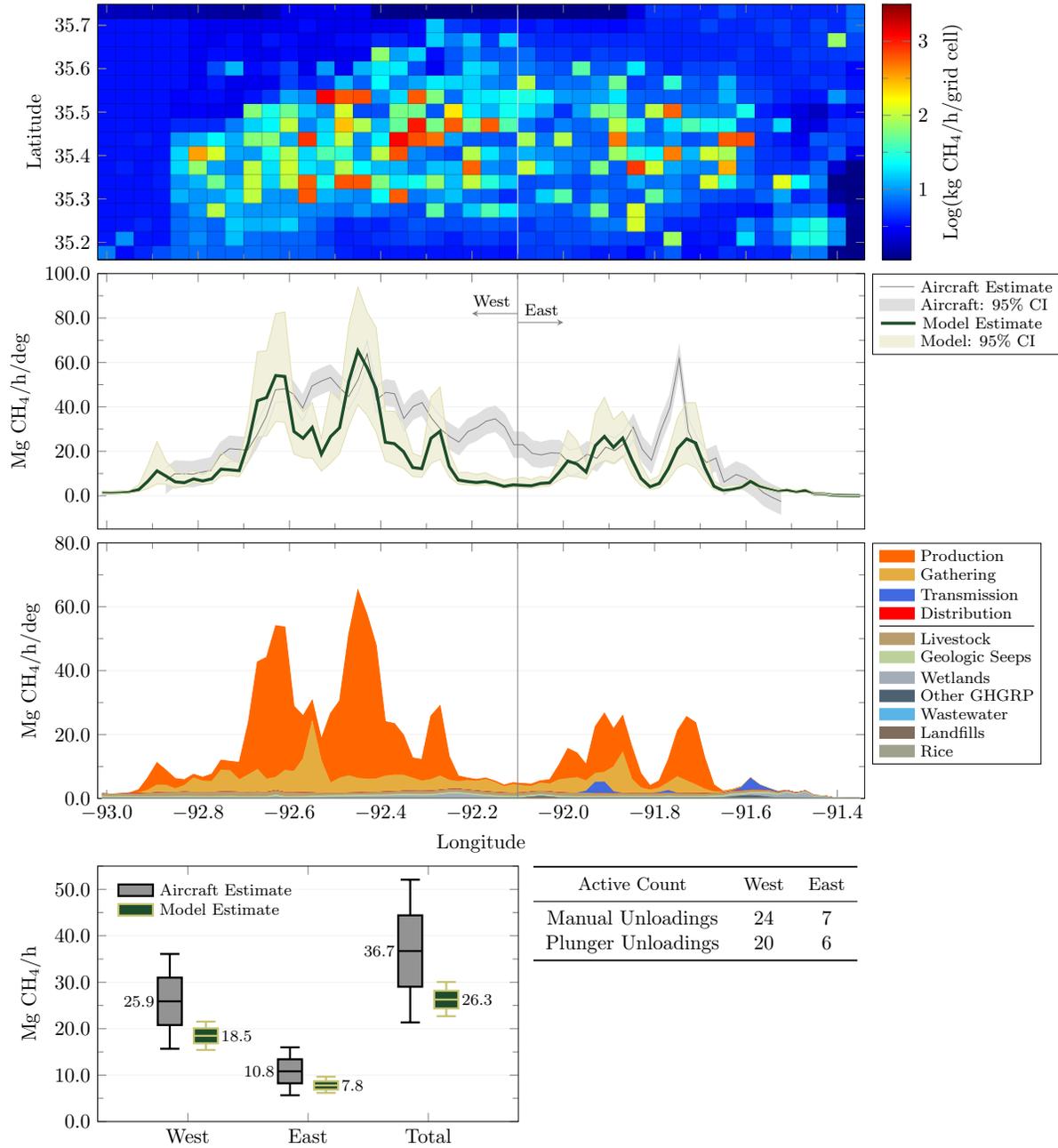
October 2, 2015 11:00-12:00 CDT



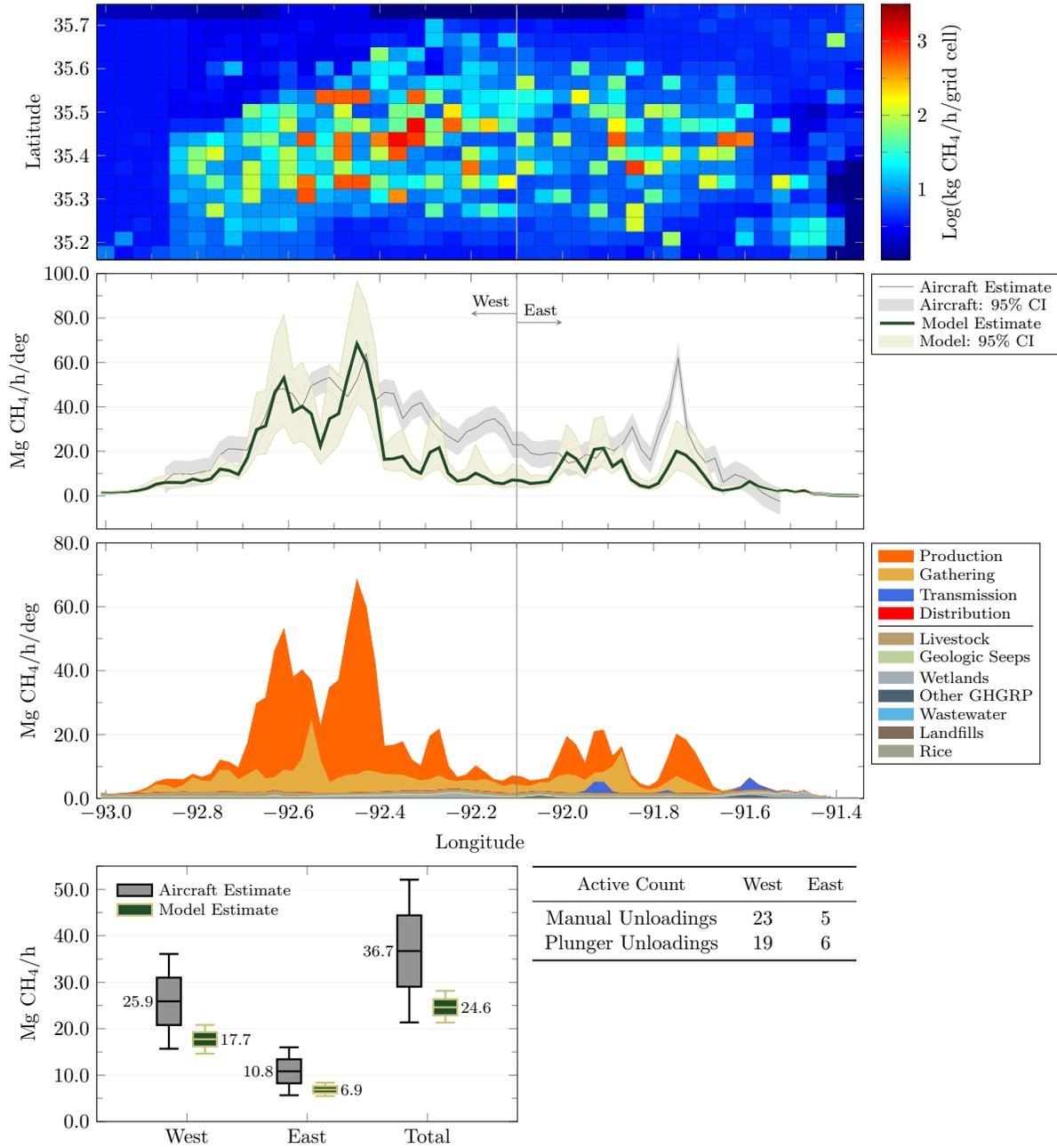
October 2, 2015 12:00-13:00 CDT



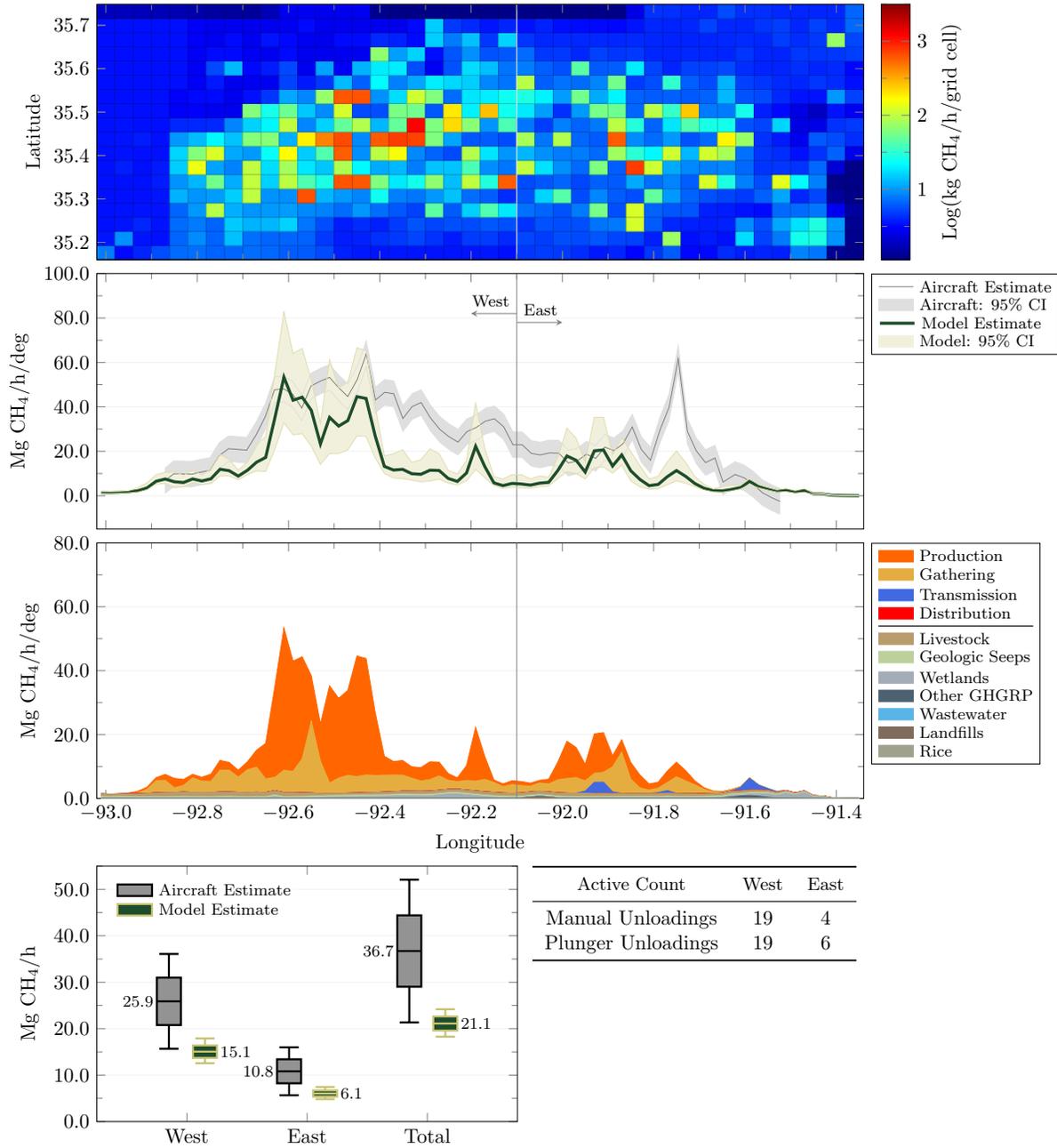
October 2, 2015 13:00-14:00 CDT



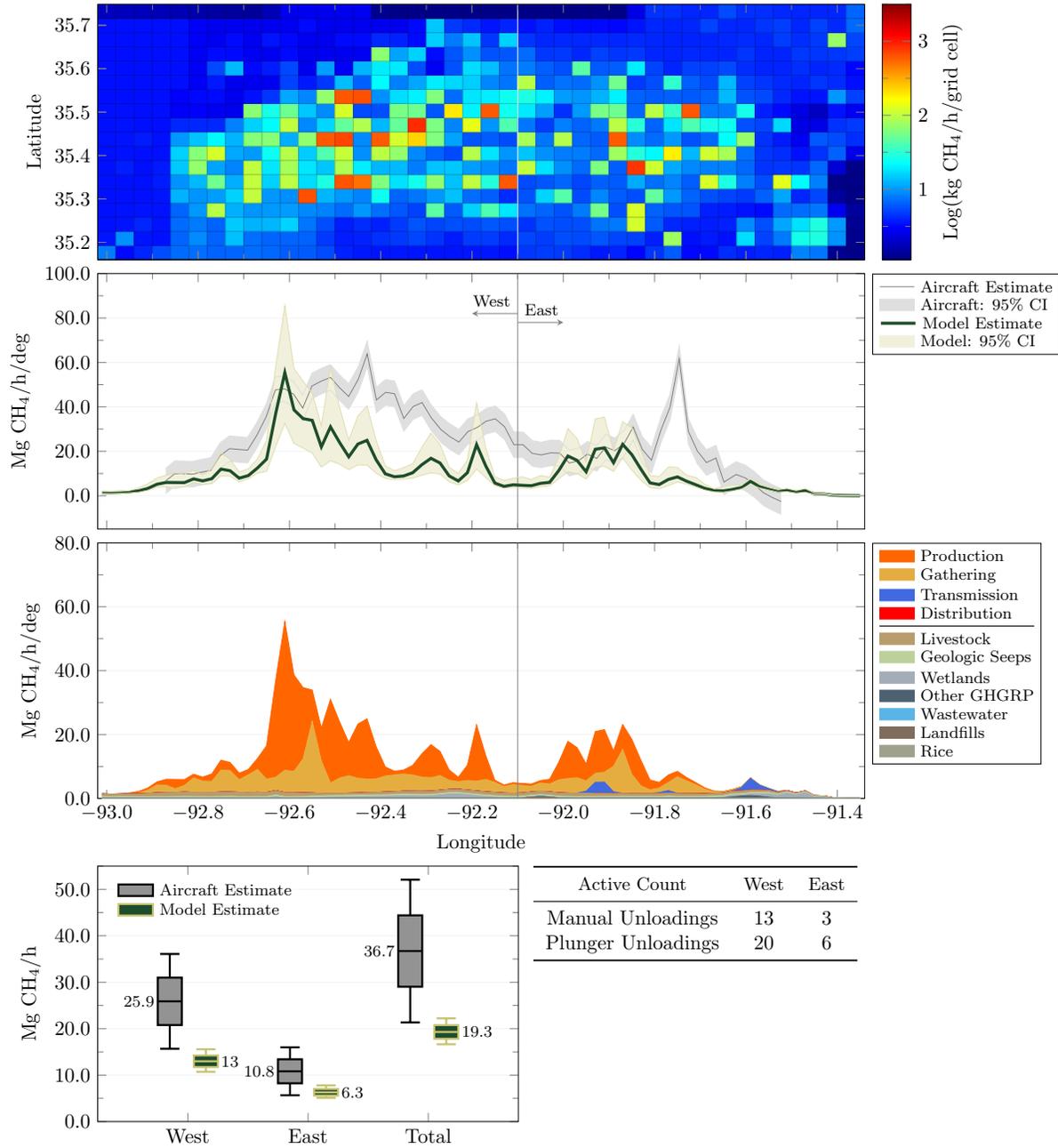
October 2, 2015 14:00-15:00 CDT



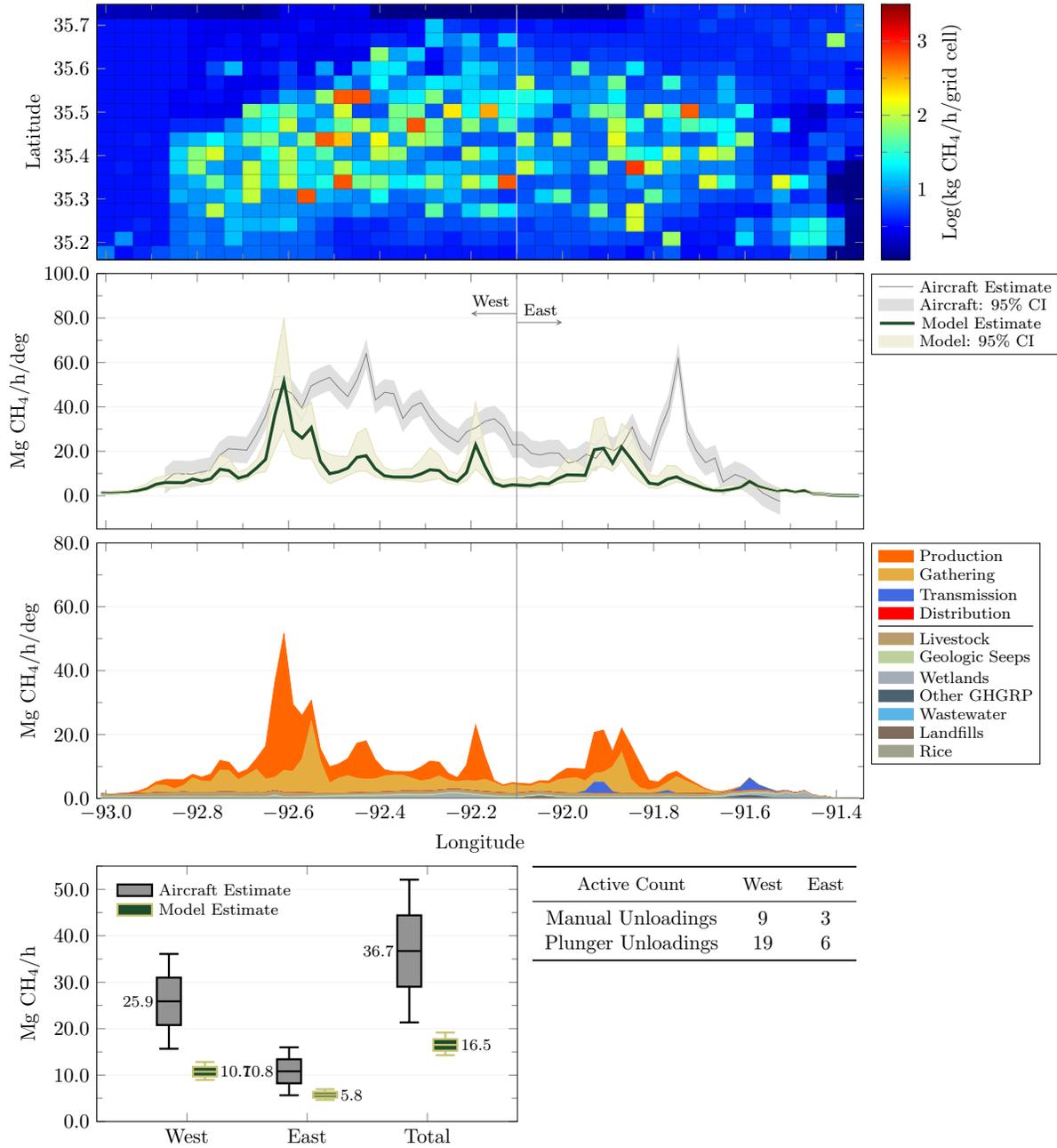
October 2, 2015 15:00-16:00 CDT



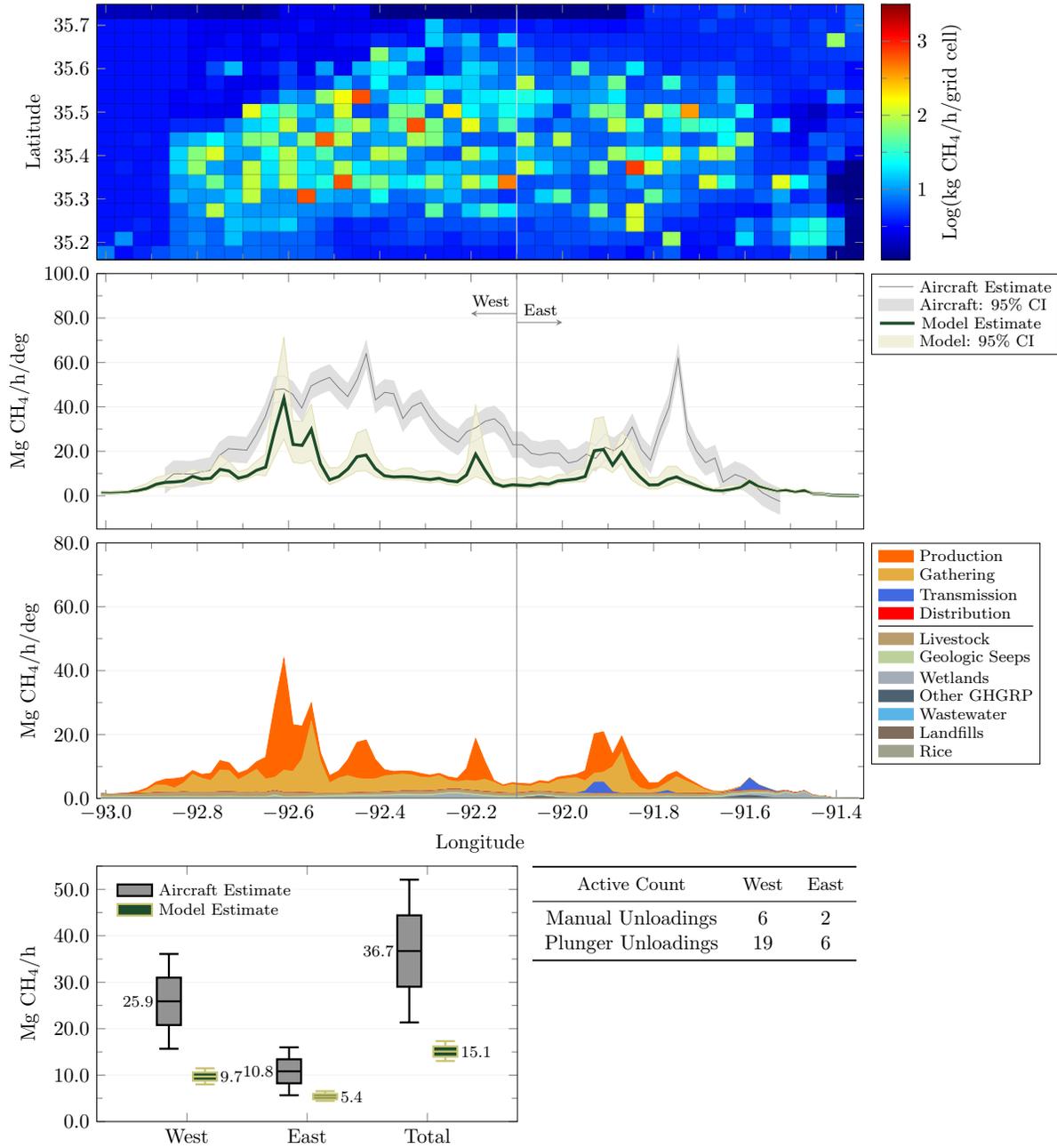
October 2, 2015 16:00-17:00 CDT



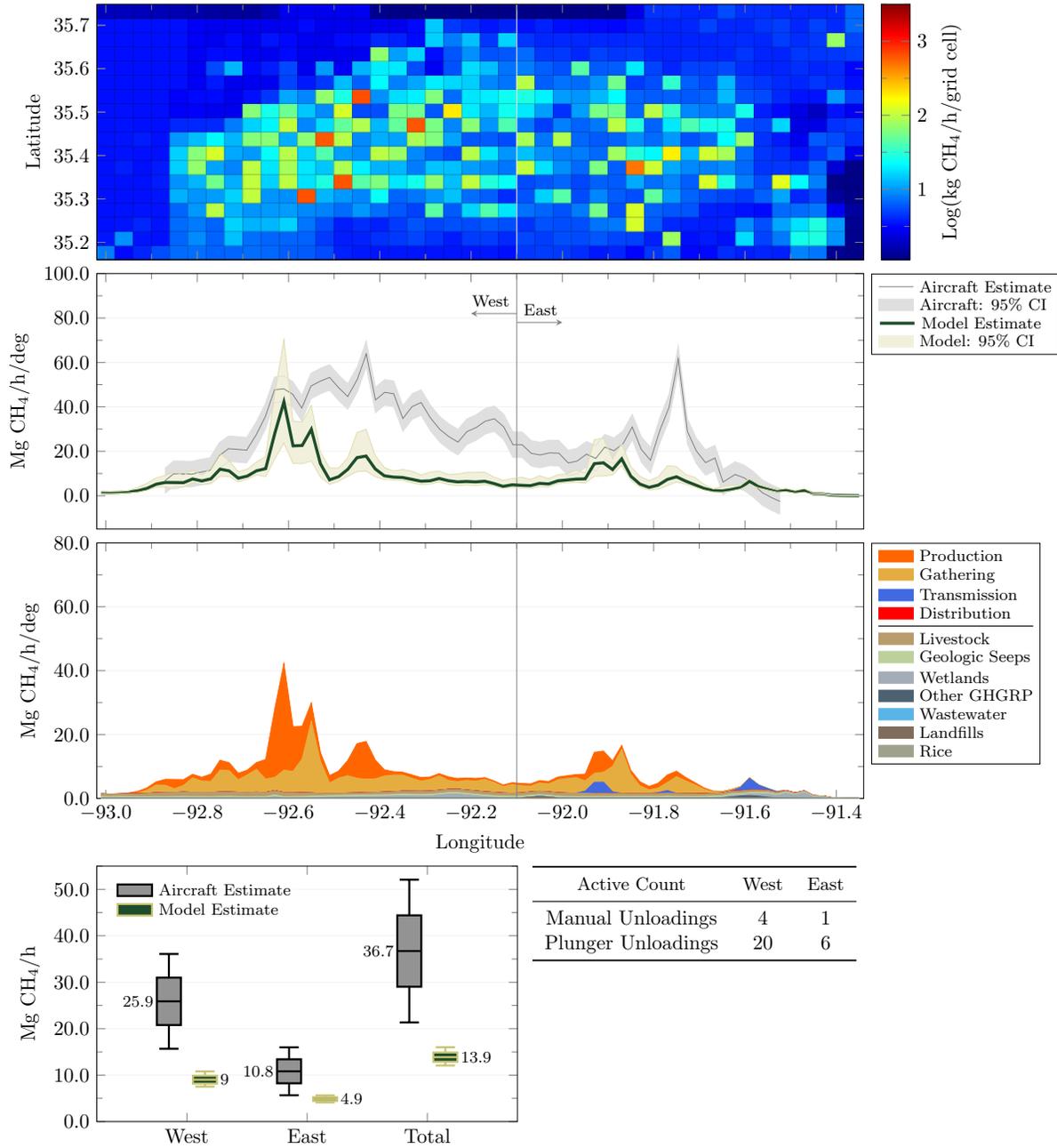
October 2, 2015 17:00-18:00 CDT



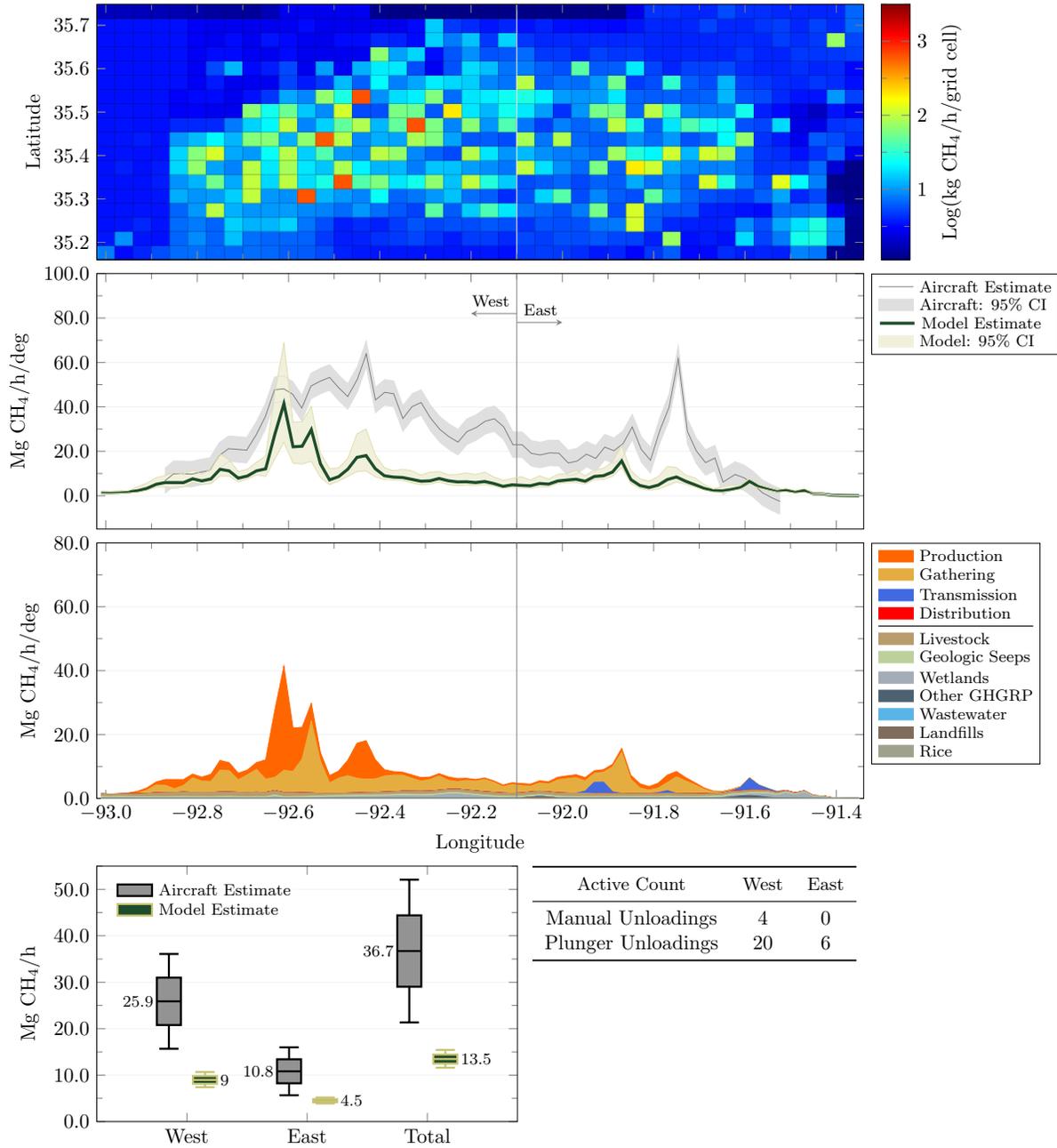
October 2, 2015 18:00-19:00 CDT



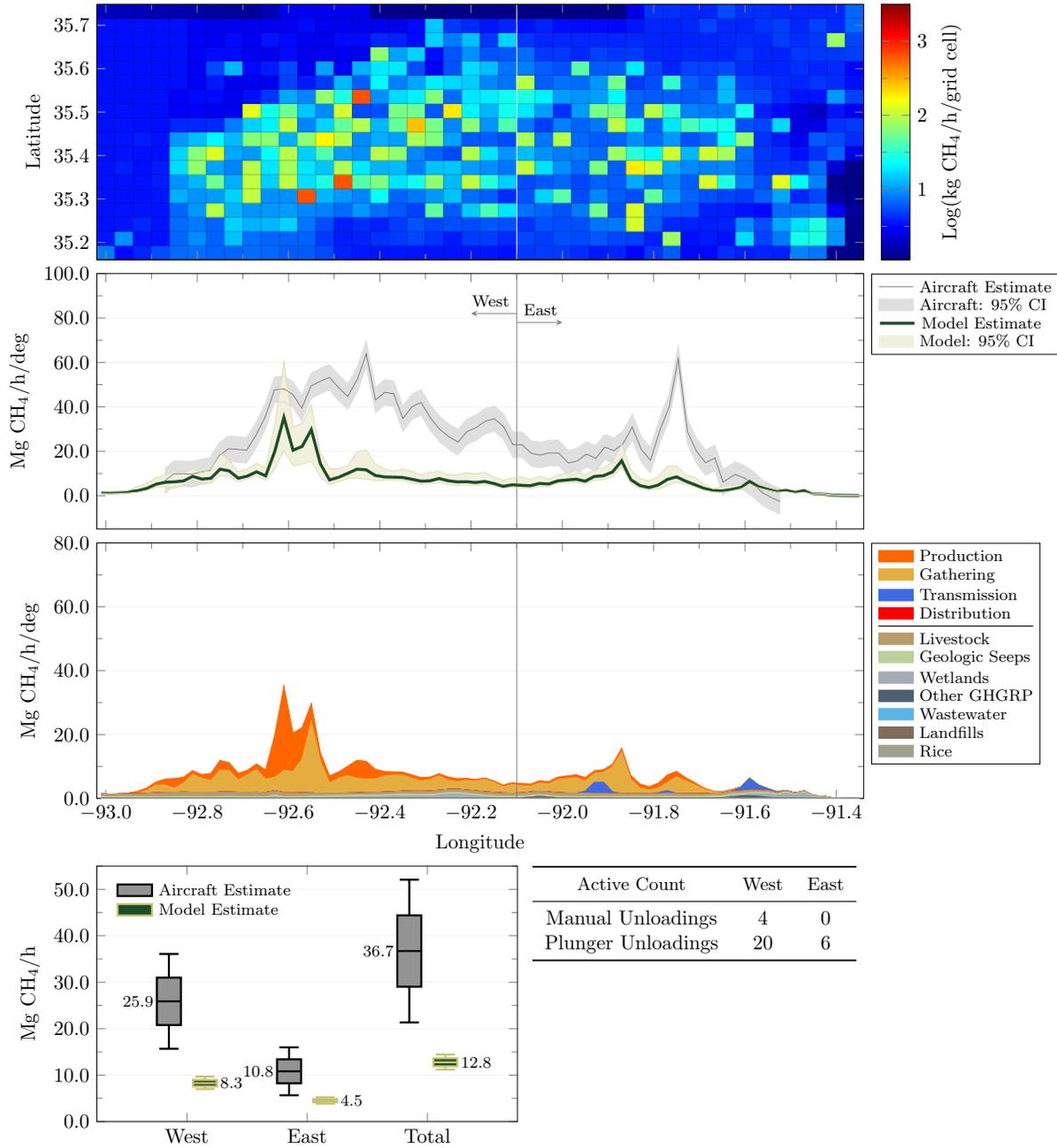
October 2, 2015 19:00-20:00 CDT



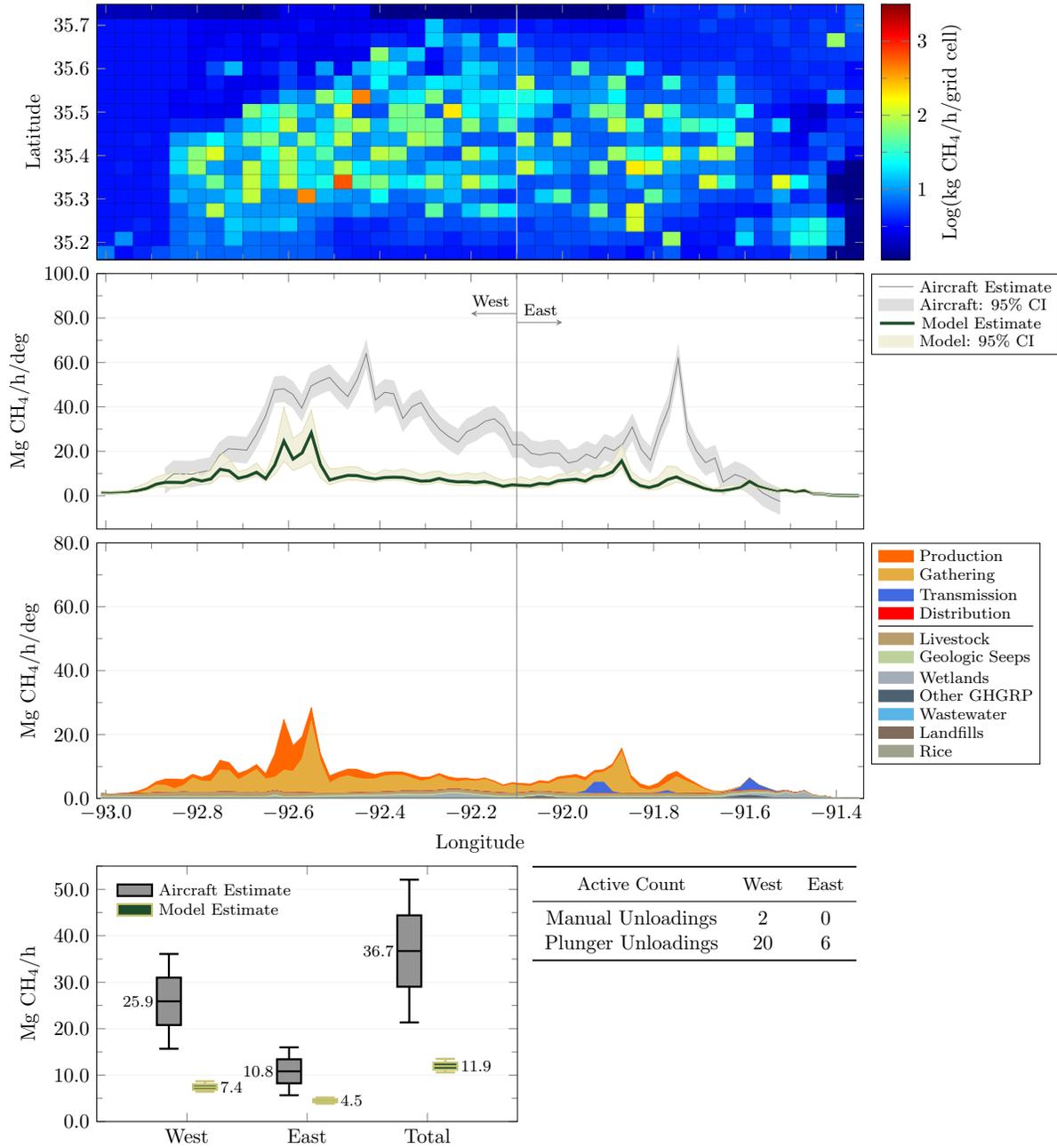
October 2, 2015 20:00-21:00 CDT



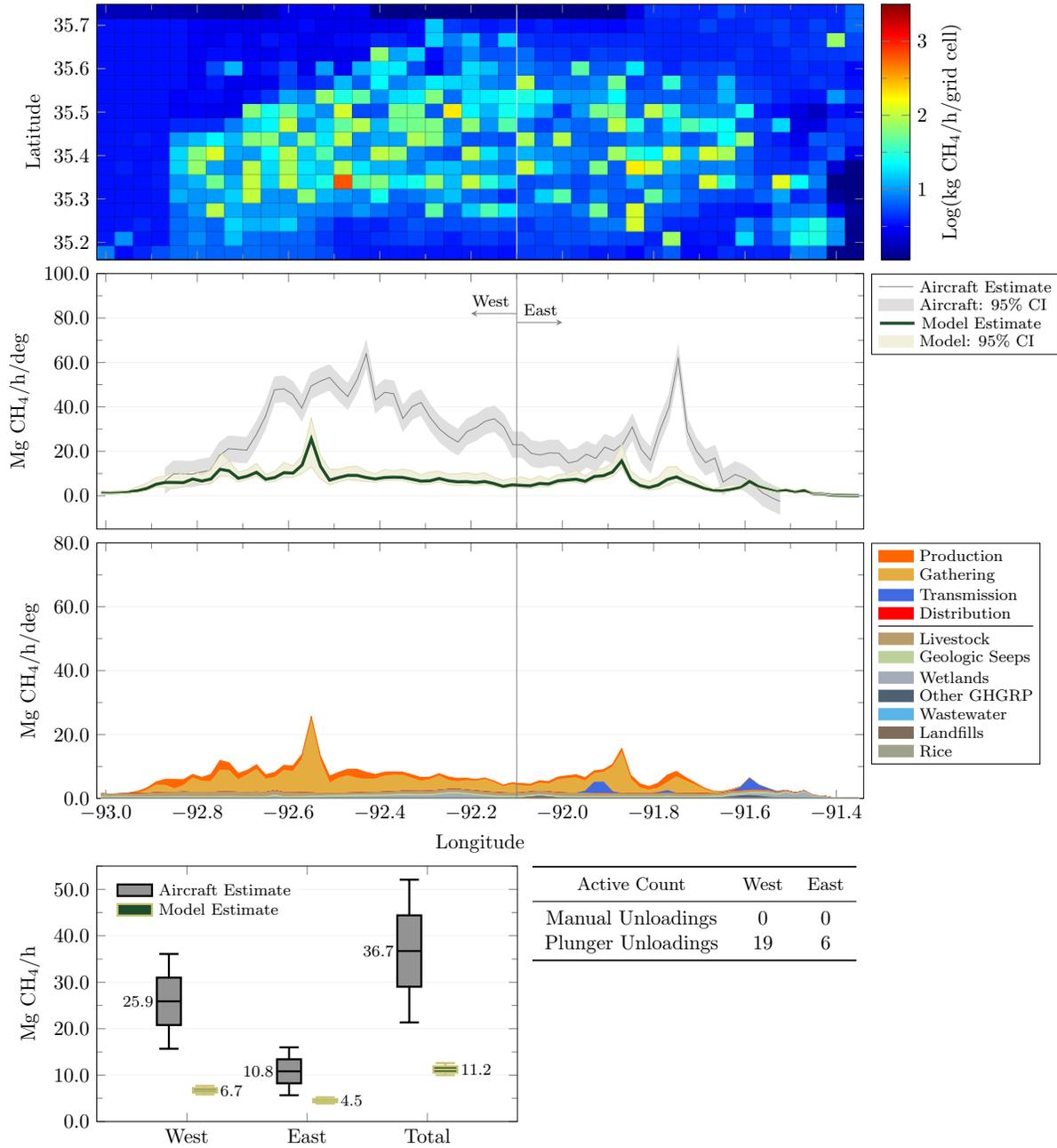
October 2, 2015 21:00-22:00 CDT



October 2, 2015 22:00-23:00 CDT



October 2, 2015 23:00-00:00 CDT



LIST OF ABBREVIATIONS

ADEQ	Arkansas Department of Environmental Quality
AFE	aircraft facility estimate
AGR	acid gas removal
AOGC	Arkansas Oil and Gas Commission
CDF	cumulative distribution function
CH₄	methane
CI	confidence interval
CO₂	carbon dioxide
CSU	Colorado State University
EIA	Energy Information Agency
EPA	Environmental Protection Agency
FLER	facility-level emission rate
FLIGHT	Facility Level Information on GreenHouse gases Tool
GHGI	greenhouse gas inventory
GHGRP	greenhouse gas reporting program
GLAE	ground-level area estimate
GRI	Gas Research Institute
GWP	global warming potential
IPCC	Intergovernmental Panel on Climate Change
LDAR	leak detection and repair
MLU	manual liquid unloading
M&R	metering and regulating
NAICS	North American Industrial Classification System
NGL	natural gas liquid
ODM	on-site direct measurement

OGI	optical gas imaging
SDM	simulated direct measurement
SIC	Standardized Industrial Classification
SOE	study on-site estimate
TCEQ	Texas Commission on Environmental Quality
TDTS	transmission distribution transfer station
TFE	tracer facility estimate
tnFLER	throughput normalized facility-level emission rate
USDA	United States Department of Agriculture
VWLS	variance-weighted least-squares