Quantifying Regional Methane Emissions in the New Mexico Permian Basin with a Comprehensive Aerial Survey

Yuanlei Chen,*∥ Evan D. Sherwin,∥ Elena S.F. Berman, Brian B. Jones, Matthew P. Gordon, Erin B. Wetherley, Eric A. Kort, and Adam R. Brandt

ABSTRACT: Limiting emissions of climate-warming methane from oil and gas (O&G) is a major opportunity for short-term climate benefits. We deploy a basin-wide airborne survey of O&G extraction and transportation activities in the New Mexico Permian Basin, spanning 35,923 km², 26,292 active wells, and over 15,000 km of natural gas pipelines using an independently validated hyperspectral methane point source detection and quantification system. The airborne survey repeatedly visited over 90% of the active wells in the survey region throughout October 2018 to January 2020, totaling approximately 98,000 well site visits. We estimate total O&G methane emissions in this area at 194 (+72/−68, 95% CI) metric tonnes per hour (t/h), or 9.4% (+3.5%/−3.3%) of gross gas production. 50% of observed emissions come from large emission sources with persistence-averaged emission rates over 308 kg/h. The fact that a large sample size is required to characterize the heavy tail of the distribution emphasizes the importance of capturing low-probability, high-consequence events through basin-wide surveys when estimating regional O&G methane emissions.

KEYWORDS: methane emissions, oil and gas, leakage, hyperspectral imaging, remote sensing, airborne survey

INTRODUCTION

Methane, the primary constituent of natural gas (NG), is a potent greenhouse gas (GHG) with a global warming potential 28−36 times larger than carbon dioxide over a 100-year time horizon and 84−87 times larger over 20 years.1 Despite the accelerating transition to renewable energy, NG continues to account for 34% of U.S. primary energy consumption as of 2020.2 Therefore, reducing the GHG intensity of oil and gas (O&G) through preventing methane emissions is an important mitigation opportunity.

The Permian Basin in Texas and New Mexico produces more oil than all but five countries in the world.3 Over the past decade, Permian oil production has quadrupled and gas production has tripled.4 However, as production from this oil-rich basin has increased, incentives to limit the resulting emissions of climate-warming methane have been lacking. Economically, operators view oil as the primary product,4 because natural gas prices in the region have remained low, or sometimes even negative, due in part to a lack of gas takeaway capacity.5 Regulations have also been slow to catch up to the pace of development - New Mexico in particular has never before had large-scale oil production, and is only now implementing state-level regulations on venting and flaring.6 Taken together, the lack of economic and regulatory incentives to reduce methane emissions has likely contributed to high methane emissions in the Permian Basin.7−9

A number of studies have found abnormally high methane emissions from O&G operations in the Permian Basin. With aircraft- and tower-based methane concentration measurements, Lyon et al. estimated the NG production loss at 3.3% in a subdomain of the Permian.8 Zhang et al. apply inversion methods based on satellite measurements by the Tropospheric Monitoring Instrument (TROPOMI), finding a NG production loss rate of roughly 3.7% for the full Texas and New Mexico Permian.7 Schneising et al. used 2018/2019 TROPOMI data and a mass estimation framework to reach a similar loss rate estimate of 3.7% in the full Permian Basin.10 Irakulis-Loitxate et al. found 37 extreme methane point sources in the Delaware sub-basin with satellites.11 These extreme sources account for 34% of regional emissions estimated by the Zhang et al. inversion method and TROPOMI measurements, revealing the heavy-tailed nature of O&G methane emissions.

Received: September 23, 2021
Revised: February 11, 2022
Accepted: February 15, 2022
Published: March 23, 2022
More recently, a hyperspectral airborne survey by Cusworth et al. characterizes the very heavy tail of site-level methane emissions in the Permian Basin, finding 2874 methane plumes above 100 kg/h and 457 above 1000 kg/h, larger than any observation previously found in ground-based methane surveys.\(^\text{12}\) Because of the different methods and coverage areas of these studies, direct comparison of their results is challenging and uncertainty remains about the emission rates in the Permian Basin.

However, these studies consistently find emissions significantly in excess of government estimates. The U.S. Environmental Protection Agency (EPA) Greenhouse Gas Inventory (GHGI) estimates a national NG production loss rate of 1.5%,\(^\text{13,14}\) but the GHGI has been identified as a conservative estimate of methane emissions,\(^\text{13,15,16}\) and a recent alternative estimate finds U.S. national average NG production loss rate of 2.3% based on a synthesis of measurements from across the O&G supply chain.\(^\text{15}\) Note that the Permian findings are even higher than this adjusted national average. One possible driver of larger emissions in the Permian might be the large point sources found by Cusworth et al.: infrequent large emissions (so-called “super-emitters”) are thought to play an important role in driving total emissions. Across many studies, the top 5% of sources contribute over 50% of emissions.\(^\text{17}\)

How are these figures still so uncertain? In short: field measurements are noisy and the high expense of surveys means that most studies to date have been very data-limited. For example: the largest multipaper synthesis data set of ground-based site-level methane measurements includes measurements from ~1000 well sites across nine different studies.\(^\text{15}\) Given that there are over one million active O&G wells in the U.S., this is a relatively small sample size. Especially given the importance of infrequent superemitters in driving total emissions, such sample sizes are difficult to extrapolate.

We bridge this gap using a novel approach: A basin-wide aerial survey capable of measuring emissions from nearly every asset in an O&G producing region with an instrument capable of quantifying and attributing medium-to-large point-source emissions. This work allows us to identify emissions larger than any documented in ground-based surveys, and to obtain sample sizes orders of magnitude larger than prior approaches.

### MATERIALS AND METHODS

#### Repeated Comprehensive Airborne Survey

We use a basin-wide data set from aerial surveys performed by Kairos Aerospace (henceforth “Kairos”) to evaluate medium-to-large point-source emissions in the New Mexico Permian Basin. Kairos’ technology consists of an integrated infrared imaging spectrometer, optical camera, global positioning system (GPS), and inertial motion unit.\(^\text{18}\) The instrument is flown on an airborne platform at ∼900 m above ground, and generates methane plume images superimposed over concurrent optical images (see example in Figure 1a). More information about this sensing technology is available in Kairos’ technical white paper and sensing systems patent.\(^\text{18,19}\) Note that the methane plume colors indicate confidence levels of methane enhancement.\(^\text{20}\)

Supporting Information (SI) Figure S1 compares this methane confidence representation with with a methane concentration enhancement representation.

Sherwin, Chen et al. evaluated the Kairos technology by conducting an independent, single-blind test of the system including 234 total measurements.\(^\text{22}\) The test found (1) no false positives among the 21 negative controls; (2) a minimum detection level of 5 kg of methane per hour per meter per second of wind (kg/h/mps) and a partial detection range of 5–15 kg/h/mps; and (3) an R\(^2\) value of 0.84 between the measured and actual release volumes across a wide range of release sizes tested (18–1025 kg/h) above the technology’s detection limit. This study showed the technology’s ability to quantify superemitters in the field.\(^\text{22}\) See the SI Section S1 for detailed controlled release results.
The Kairos survey of the New Mexico Permian was conducted over 115 flight days from October 2018 to January 2020 (Figure 1b). The campaign surveyed 35,923 km² (13,870 sq. mi.) and 26,292 active wells, or 91.2% of all active wells in the covered region. All data were anonymized using procedures described in SI Section S2.2.

Each surveyed nonpipeline facility was observed an average of four times. Accounting for these repeated measurements, the Kairos survey performed a total of 117,658 visits to wells, or approximately 98,000 well site visits based on 1.2 wells per well site in the New Mexico Permian Basin found by a 2018 ground survey.9

Figure 1c shows the number of measurements of each point asset (nonpipeline). Multiple overflights also allowed for more frequent sampling in the temporal dimension and provided insights into emission intermittency. The SI Section S2 details the flight plans and SI Section S3 presents an analysis of intermittency.

**Basin-Wide Emissions Quantification.** A methane survey will detect some number of plumes, each of which is associated with an emission source. An emission source is defined as a point coordinate with one or more methane plumes observed during the campaign. SI Section S4.2 describes the plume-source association process.

Figure 2 illustrates the analysis workflow to derive survey-area total emissions. SI Section S5.1 describes each step in detail. For each plume, Kairos reports a wind-independent emission rate in kg/h/imp, and we multiply this rate with the National Oceanographic and Atmospheric Administration’s High Resolution Rapid Refresh (HRRR) wind speed estimate at the imaging time and plume coordinates to calculate emission rate in kg/h for each plume using the method described in Duren et al.23

We then refer to the single-blind test of the instrument by Sherwin, Chen et al. to determine the instrument’s detection limit and quantification accuracy and precision (see SI Section S1). Data from the single-blind test show an apparent overestimation tendency when simply multiplying HRRR wind with the reported wind-independent emission rates for larger releases, possibly due to an underlying nonlinearity or a boundary bias for calibration (see SI Section S1.6). Using a power law correlation from the single-blind test, we calibrate the plume-level emission rates in kg/h. This correlation (detailed in SI Section S1.5) corrects for the apparent overestimation tendency for large releases when using HRRR wind. The single-blind test also quantified the measurement uncertainties, which are modeled as a fixed percent error distribution at all emission levels, indicating that the modeled absolute error scales linearly with emission magnitude (see SI Section S1.5). To account for the measurement error in the New Mexico Permian Basin study, we assume that the percent error follows a normal distribution and apply this error to the plume-level emission rates with 1000 Monte Carlo realizations.

For each realization of the Monte Carlo approach, we then select one plume for each emission source if multiple plumes were observed during repeated overflights. Then we multiply the selected plume quantification with a binary term to account for intermittency. Each emission source has a probability, p, of emitting in a given Monte Carlo iteration, with p equal to the fraction of overflights that observed emissions at each emission source.

Basin-wide directly measured emissions \(\hat{E}_{\text{measured}}\) is the sum of all emission source level emissions after accounting for intermittency. Note that we include all emission observations, including those from emission sources that were covered only once in the campaign for basin-wide emission quantification. Although one observation is not sufficient to characterize the time-averaged emission rate of a single source, a basin-wide survey measuring a large number of sources one time (or multiple times in the case of this study) is sufficient to give an unbiased estimate of the whole basin. See SI Section S3 for further detail.

For simplicity, we assume that the distribution of methane emissions is stationary over time, although we observe some evidence of seasonal and intraday variation in the frequency of aerially visible methane emissions. The direction of the effect of this variation on our estimate is unclear. SI Section S3.6 explains why \(\hat{E}_{\text{measured}}\) is an unbiased estimate of total measured emissions.

To account for undetected emissions in the partial detection range of Kairos’ technology, we add to \(\hat{E}_{\text{measured}}\) the expected amount of emissions undetected within the partial detection range based on both the detection probabilities and what was observed in the partial detection range during the New Mexico Permian campaign (see SI Sections S1 and S5.1). We then scale up the estimate to the full study area, the black polygon in Figure 1b, assuming that emissions in uncovered areas scale with the number of O&G wells in the area.

Below Kairos’ minimum detection threshold, we assume that emissions are described by a combination of the fractional loss rate from Alvarez et al. of 2.2% for production and midstream as well as the emission size distribution from Omara et al.13,15 Assuming winds from the New Mexico Permian, Kairos would be able to detect 63% of emissions from Omara et al. 2018, translating to a fractional loss rate of 0.8% for emissions below the detection threshold in this study. See SI Sections S1.4 and S5.1 for partial detection definition and detailed steps to account for undetected emissions. We denote the total emissions after incorporating undetected emissions as \(\hat{E}_{\text{total}}\).

### RESULTS AND DISCUSSION

**Large Basin-Wide Methane Emissions Quantified.** The campaign detected 1985 methane plume observations from 958 distinct emission sources, indicating that for the average emissions source, approximately two different overflights observed a plume. Using the approach described in the Materials and Methods, our estimate for measured emissions \(\hat{E}_{\text{measured}}\) from the New Mexico Permian is 153 (+71/−70, 95% CI) metric tonnes per hour (t/h), shown as the left bar in
Figure 3. Persistence-averaged emissions. (a) The left bar shows directly measured methane emissions (E_{measured}) broken down by asset type. The error bars indicate 95% confidence intervals. The middle bar breaks down extrapolated emissions into undetected emissions within the partial detection range (PDR), emissions from assets not measured in the survey area, and emissions that are below minimum detection limit (MDL). The right bar shows that the estimate of total methane emissions in the survey area from upstream and midstream O&G operations is 194 (+72/−68) t/h, 9.4% (+3.5%/−3.3%) of gross gas production. (b) The distribution of asset-type-specific persistence-averaged emission source sizes, which follow heavy-tailed distributions. (c) Cumulative emission fraction as a function of persistence-averaged emission source sizes.

Figure 3a. This corresponds to 7.4% ± 3.4% of gross gas production in the full survey area.

Accounting for partial detection, emissions below minimum detection limit, and scaling up to assets not covered in this aerial campaign, the total survey area emission estimate (E_{total}) is 194 (+72/−68) t/h, equivalent to 9.4% (+3.5%/−3.3%) of gross gas production.

A breakdown of E_{measured} by emission source asset type reveals that 79 ± 46 of the 153 t/h of measured emissions comes from well sites. A "well site" is defined here as the ensemble of all assets (including wells, gathering lines, storage tanks, and compressor stations) found on a congruent gravel or concrete area containing at least one well. Midstream assets were also a significant source, with 29 ± 20 t/h emitted from pipelines (including underground gas gathering pipelines) and 26 ± 16 t/h emitted from compressor stations without a well on site. The remainder was emitted from stand-alone storage tank sites (9 ± 6 t/h), gas processing plants (4 ± 2 t/h), and other or ambiguous sources (7 ± 4 t/h). See SI Section S4.2 for definitions of each asset type and the asset attribution method.

Figure 3b shows the distribution of persistence-averaged emission source sizes and indicates heavy-tailed distributions of emission sizes across asset types. As displayed in Figure 3c, 50% of measured emissions are from 118 (~12%) of the 958 sources, those larger than 308 kg/h. The heavy tail gets even heavier for the largest emissions and contains a disproportionate number of midstream assets. The largest persistence-averaged emission source emits at 4.3 t/h. The persistence of the heavy tail for distributions of large emissions demonstrates the significant potential for mitigating methane by detecting and fixing these high-consequence sources.

Sensitivity tests show robust support for a mean natural gas fractional loss rate of at least 8.1% of gas produced. As listed in Table 1, switching from a power law fit to a linear fit for the calibration step, described in SI Section S7, brings the loss rate estimate up to 10.2% (+4.1%/−3.6%). A linear fit forced through the origin leads to an estimate of 11.0% (+5.0%/−4.6%). In the calibration fitting process, leaving out large controlled releases improves the statistical validity of the fit due to the underlying asymmetric error distribution at high emission rates, and also increases the total emission estimate, as described in the SI Section S1.5. Using an alternative wind data set (the commercial Dark Sky wind reanalysis product) results in comparable emissions estimates both for low- and high-time-resolution versions of the data.²⁴

<table>
<thead>
<tr>
<th>cases</th>
<th>E_{total} (t/h)</th>
<th>%NG production loss</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>5th%</td>
</tr>
<tr>
<td>base case</td>
<td>194</td>
<td>126</td>
</tr>
<tr>
<td>linear fit for calibration</td>
<td>212</td>
<td>136</td>
</tr>
<tr>
<td>linear fit forced through origin for calibration</td>
<td>228</td>
<td>131</td>
</tr>
<tr>
<td>cutoff at 1σ below max controlled release</td>
<td>216</td>
<td>137</td>
</tr>
<tr>
<td>dark sky wind high time resolution</td>
<td>181</td>
<td>124</td>
</tr>
<tr>
<td>dark sky wind low time resolution</td>
<td>217</td>
<td>142</td>
</tr>
<tr>
<td>disable extrapolation</td>
<td>167</td>
<td>119</td>
</tr>
<tr>
<td>exclude top 20 plumes</td>
<td>173</td>
<td>117</td>
</tr>
<tr>
<td>no emissions below minimum detection</td>
<td>177</td>
<td>109</td>
</tr>
</tbody>
</table>

To provide a conservative estimate for the loss rate, we apply three additional sensitivity scenarios: (1) disallow extrapolation and assume that emission rates cannot exceed the largest controlled release rate (1025 kg/h); (2) exclude the top 20 largest plumes (~1% of the data set); and (3) assume that there are no emissions from plumes below the Kairos minimum detection limit. These conservative approaches still result in mean loss rate estimates over 8% with a 5th percentile estimate never falling below 5.2%.

These sensitivity cases show that even the lower-bound estimates of the conservative scenarios based on our basinwide data are larger than estimates from other Permian studies: 3.7% by the Zhang et al. and Schneising et al. satellite-based top-down studies and 3.3% by the Lyon et al. tower- and airplane-based top-down study, although these studies include
both Texas and New Mexico. Applying our basin-wide quantification method to data from Cusworth et al. in the overlapping region of New Mexico, we find a fractional loss rate of 4.4% for directly measured emissions. This rises to 5.9% after accounting for an evidently higher effective minimum detection threshold compared to the Kairos survey (see SI Section S8).

The reasons for these discrepancies are currently unknown. Increasing evidence suggests that strong time trends exist in Permian flaring and emissions, and that 2019 was a period of rapid production growth, large amounts of flaring, and presumably poor gas management in general. If this is the case, then our study period could have higher actual loss rates than other study periods.

More work will be required to understand why our results do not align with satellite-based top-down studies. It is important to note that our study is based upon blind-validated methods using hundreds of third-party validation measurements (as seen in SI Section S1). We believe that comprehensive regional aerial surveys with single-blind-validated instruments could provide an empirical basis for calibrating such top-down models, which has historically been difficult due to the large modeling scale.

**Importance of Large Sample Size and Direct Measurement.** Compared to an EPA GHGI estimate aligned to our study area and time period (Figure 4a), this study suggests total methane emissions from upstream and midstream O&G activities in the New Mexico Permian to be 6.5 (±2.4/−2.3) times larger. It is important to explore further a key strength of our method compared to prior bottom-up studies: very large study sample size. We explore this by simulating the impact of small sample sizes on total emissions estimates (Figure 4b).

Suppose that we only visited 100 well sites, a typical sample size for ground-based campaigns. Based on a random subsample of 100 well site visits from our full data set of 98 000 effective well site visits, and using the same minimum detection limit as Kairos, this hypothetical 100-well site survey would detect no emissions 34% of the time and would find average emissions lower than the basin-wide survey 72% of the time (based on 1000 Monte Carlo realizations). Median emissions would be 34% of our full survey estimate. In a small number of Monte Carlo realizations (12%), scaling up the 100 sampled visits results in overestimates by a factor of 2 or more. Over many Monte Carlo realizations, a sample size of 100 will ultimately converge on the larger survey results, but this does not reflect the reality of field campaigns: there are usually no more than a few such campaigns for a given basin in a given decade and averaging over 1000 hypothetical surveys does not apply.

Figure 4b shows that increasing the sample size per simulated survey to 1000 well site visits generates an underestimate of total emissions 63% of the time, while a size of 10 000 effectively captures large-scale behavior. The extremely non-normal distribution of emission sizes plays a large role here and intuition developed with normally distributed phenomena may be deceiving. In normally distributed phenomena, small sample sizes cause variance but not bias, and increasing sample size reduces the variance in the estimated emissions. But with our observed contribution of superemitters, the median estimate of a simulated survey shifts strongly to the right as our sample size increases: at 100 well site visits the median estimate is 42% of our estimate, at 1000 visits this increases to 82%, and at 10 000 visits it increases to 99% of our estimate.

**Airplane-Detectable Emitters Drive Total Emissions.** While aerial detection technologies have been critiqued for their relatively high minimum detection limit, our results suggest an alternative interpretation: the error introduced from the small sample sizes feasible with ground campaigns may overwhelm any benefits they get from a lower detection threshold. For example, below-minimum-detection-limit emissions account for 9% (±4%−3%) of our study total, suggesting that higher sensitivity would lead to only a modest
increase in total estimated emissions relative to simulated levels. In conclusion, we conducted a site-level, basin-wide field survey of methane emissions in one of the most active oil-producing regions in the world. We estimate emissions to be 9.4% (+3.5%−3.3%) of the gross gas production for the region, much higher than found in previous studies with overlapping, although not identical, domains. The increase is partly because our method allows us to inspect the entire O&G-producing population using an independently verified instrument capable of detecting large methane emissions. This allows us to identify the largest emissions from all assets surveyed, sidestepping the statistical uncertainties of scaling-up small samples of ground-based field measurements.

Previous studies rarely observed emissions larger than 10 kg/h at a single site, yet our basin-wide survey of over 30 000 assets uncovered 1958 methane plumes above this size.9,15 This includes many emissions over 100 and 1000 kg/h, with emissions above 308 kg/h accounting for half of measured emissions for the region. While it is possible that the New Mexico Permian was an anomaly during this study period, the clear impact of large emissions found by this study suggests that estimates from ground-based methane surveys may be underestimating total emissions by missing low-frequency, high-impact large emissions.

**ASSOCIATED CONTENT**

**Supporting Information**
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.1c06458.

Details of the technology, the New Mexico Permian survey, intermittency analysis, emission attribution, quantification method, sensitivity cases, and comparison with other Permian studies (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**
Yuanlei Chen — Energy Resources Engineering, Stanford University, Stanford, California 94305, United States; orcid.org/0000-0002-4341-2414; Email: yuliachen@stanford.edu

**Authors**
Evan D. Sherwin — Energy Resources Engineering, Stanford University, Stanford, California 94305, United States; orcid.org/0000-0003-2180-4297
Elena S.F. Berman — Kairos Aerospace, Mountain View, California 94040, United States
Brian B. Jones — Kairos Aerospace, Mountain View, California 94040, United States
Matthew P. Gordon — Kairos Aerospace, Mountain View, California 94040, United States
Erin B. Wetherley — Kairos Aerospace, Mountain View, California 94040, United States
Eric A. Kort — Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, Michigan 48109, United States
Adam R. Brandt — Energy Resources Engineering, Stanford University, Stanford, California 94305, United States; orcid.org/0000-0002-2528-1473

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.1c06458

**Author Contributions**
Y.C. and E.D.S. contributed equally to this work.

**Notes**
The authors declare the following competing financial interest(s): E.S.F.B., B.B.J., M.P.G., and E.B.W. are employees of Kairos Aerospace. The remaining authors have no competing interests to declare.

**ACKNOWLEDGMENTS**

We thank the Kairos Aerospace team for collecting and preparing the data for this study. We gratefully acknowledge the help from Ritesh Gautam, Ben Hmiel, David Lyon, and Mark Omara at the Environmental Defense Fund, Yuzhong Zhang currently at Westlake University, Anna Robertson and Shane Murphy at the University of Wyoming, and Daniel Cusworth and Riley Duren at Carbon Mapper, and Andrew Thorpe at NASA’s Jet Propulsion Lab for assisting in comparing this aerial survey with their methane studies. We thank Jeffrey Rutherford at Stanford University for providing comments on the study. This study was funded by the Stanford Natural Gas Initiative, an industry consortium that supports independent research at Stanford University. Analysis was supported in part by the Alfred P. Sloan Foundation Grant G-2019-12451 in support of the Flaring and Fossil Fuels: Uncovering Emissions and Losses (F3UEL) project.

**REFERENCES**

(10) Schneising, O.; Buchwitz, M.; Reuter, M.; Vanselov, S.; Bovensmann, H.; Burrows, J. P. Remote sensing of methane leakage...


**EDITOR’S NOTE**

The data required to reproduce key results in this article are not available for open release due to confidentiality concerns, Kairos Aerospace is committed to working with research groups studying methane emissions. Access may be granted, but must be done directly through Kairos Aerospace. Interested researchers should contact research-collaborations@kairosaerospace.com.