Investigating aerial data pre-analysis schemes and site-level methane emission aggregation methods at LNG facilities

Olga Khaliukova,[*,†] Yuanrui Zhu,[‡] William S. Daniels,[†] Arvind P. Ravikumar,[‡,¶] Gregory B. Ross,[§] Selina A. Roman-White,[§] Fiji C. George,[§] and Dorit M. Hammerling[†,¶]

†*Department of Applied Mathematics and Statistics, Colorado School of Mines, Golden, Colorado, United States*

‡*Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, Texas, United States*

¶*Energy Emissions Modeling and Data Lab, The University of Texas at Austin, Austin, Texas, United States*

§*Cheniere Energy Inc., Houston, Texas, United States*

E-mail: okhaliukova@mines.edu

Abstract

Methane measurements at liquefied natural gas (LNG) facilities play an important role in characterizing methane emissions from the natural gas supply chain. The large size and complexity of LNG facilities make quantifying emissions with ground-based monitoring systems challenging, making aerial platforms one of the preferred methods for methane measurements at these sites. However, aerial measurements typically provide a snapshot of emissions at a given instance, necessitating further analytical steps to

infer both annualized emissions and the range of possible emissions at different instants in time. This study uses aerial measurements at two LNG facilities from a Quantification, Monitoring, Reporting, and Verification project to characterize the distribution of temporally averaged emissions (i.e., annualized inventories) and possible site-level emissions at any given point in time ("instantaneous emissions"). The former provides uncertainty on the aerial measurement component of measurement-informed inventories representing annual averages, while the latter will help contextualize future snapshot measurements (e.g., aerial surveys or high-resolution satellite platforms) at LNG facilities. We find that instantaneous emissions may fall well outside of the distribution describing uncertainty in the annual inventory. We also compare different pre-analysis schemes for aerial data, as existing literature does not provide a clear consensus on methods for doing so, especially at LNG facilities.

Keywords: methane emissions, site-level emissions, instantaneous emissions, oil and gas, LNG, measurement-informed inventory

Synopsis: We evaluate the impact of initial aerial data pre-analysis schemes and highlight the differences in distributions between instantaneous and average-based site-level total emission rates at two LNG facilities.

1 Introduction

Natural gas is often considered a transitional energy source,[1] serving as a bridge from coal and oil to renewable energy sources. Since natural gas is predominantly composed of methane (CH_4),[2] there has been increasing research into the impact of methane emissions from the natural gas value chain.[3–9] Methane is a potent greenhouse gas (GHG) with a global warming potential 82.5 times greater than that of carbon dioxide (CO_2) over a 20-year period.[10] Thus reducing methane emissions from the oil and gas industry is a crucial step in addressing global warming. This process is relatively straightforward compared to other methane emission sources such as agriculture and landfills,[11] due to the low cost of abatement, identifiable

and concentrated point sources, opportunities for equipment repair, and the availability of existing infrastructure and technologies.[12]

The Russian invasion of Ukraine has shifted geopolitical dynamics, prompting the European energy market to seek alternative energy sources to Russian pipeline gas while maintaining its strong commitment to reducing greenhouse gas emissions.[13] Concurrently, the United States has substantially increased its liquefied natural gas (LNG) exports, becoming the world's largest exporter in 2023.[14] This surge highlights the critical need for accurate methane emissions reporting at LNG facilities, which is essential for effective emission reduction strategies. However, methane emissions during the liquefaction process, which involves cooling and condensing natural gas into a liquid form for efficient storage and transportation over long distances, have not been studied as extensively[15] as those at simpler oil and gas production sites[16] or compressor stations.[17]

Thanks to voluntary programs such as the Oil and Gas Methane Partnership 2.0[18] and differentiated gas,[19] many operators in the United States are increasingly committed to developing precise, site-level measurement-based methane emissions inventories that include an aerial measurement component, emissions below aerial technology detection limits, and known emissions that are typically not captured by the overflights (e.g., Refrigeration Turbines, Thermal oxidizers). However, creating measurement-informed inventories (MII) at LNG facilities presents unique challenges due to their complexity, large scale, tall multi-level structures, and numerous equipment groups.

Various technologies and platforms are used to measure methane emissions[20] and ultimately develop measurement-informed methane inventories at various scales.[21–23] Many of these technologies have been tested via controlled releases.[24–26] However, a small number of measurements taken at a single instant in time ("snapshot measurements") can under- or overestimate long term emission characteristics due to the temporal variability of emissions.[7] Measurements at a high temporal resolution from continuous monitoring systems (CMS) can be used to reconcile long term averages (i.e., inventories) and snapshot measurements at the

site-level,[27] but current CMS solutions are mostly intended for use at simplier production sites.[28,29] Therefore, to measure methane over the large and complex LNG facilities in this study, we use data from Bridger Photonics Gas Mapping LiDAR.[30] To account for temporal variability of the emissions, Bridger conducted multiple measurement campaigns at both LNG facilities, with each campaign consisting of one or more full facility scans.

Estimated methane emission rates from Bridger have been evaluated in a number of single-blinded studies.[31–35] However, quantifying emission rates poses a challenge when the detected plume is located at the edge of Bridger's measurement swath, a region approximately 100 meters wide captured by the optical sensor. When this occurs, Bridger typically notes that an emission has occurred, but does not produce a rate estimate. Currently, the literature does not contain clear guidelines on how to use these measurements. In this paper, we explore three data pre-analysis schemes, one of which employs a similar approach to that used by Johnson et al.[36] We demonstrate how these schemes influence the resulting distribution of point-source emissions.

Once pre-analysis schemes are finalized, additional analytical choices remain for aggregating individual measurements into a comprehensive site-level inventory. Other studies have focused on evaluating greenhouse gas emissions across the life cycle of LNG but lack details on the computational methods for LNG site inventories. Schuller et al.[37] and Okamura et al.[38] reported emissions associated with LNG at the producing country level, but provided limited detail on equipment- or site-level emissions. Tamura et al.[39] studied the life cycle of CO_2 and CH_4 emissions in LNG, using survey data to track emissions during the liquefaction process, but do not describe the methods for computing total site emissions. Zhang et al.[40] predominantly focused on CO_2 emissions at LNG facilities and emphasized that a lower detection limit is necessary for their method to effectively measure methane emissions. Innocenti et al.[15] quantified emission rates for key functional elements (site zones corresponding to particular processes) at LNG facilities using remote sensing techniques but highlighted the need for more direct measurements to produce annual site-level rates. Building on these

studies, Zhu et al.[41] evaluated the first multi-scale measurements of GHG emissions at two US liquefaction terminals and found that the measured GHG emissions intensity is consistent with the reported emission intensities, as CO_2 contributes more than 90% of total GHG emissions on these sites.

There is no "traditional" method for establishing a MII at LNG sites, and various forms of emissions averaging can be employed when inferring site-level distributions for emission inventories. Our study adds to the work of Zhu et al.[41] by exploring two distinct approaches for computing site-level LNG emissions. We introduce a method for creating a distribution of site-level methane emissions at any given instance in time, referred to as an instantaneous emissions profile. This distribution of instantaneous emissions shows the range of emissions that a remote sensing technology, such as a satellite or aerial flyover, could detect at any given time, assuming the measurement campaign accurately captures all possible emission states at the point source-level. We compare the distribution of instantaneous emissions to a method that relies on spatial aggregation and temporal averaging, similar to Schwietzke et al.,[42] who investigated the discrepancies between episodic emission sources and traditional reporting methods at a basin level. To the best of our knowledge, this analysis utilizes the most comprehensive aerial dataset currently available for LNG sites.

2 Methods

2.1 Design of aerial measurement campaigns

This study uses methane emissions data collected by an aerial platform during the LNG Quantification, Monitoring, Reporting, and Verification (QMRV[43]) project. Specifically, Bridger Photonics provided data at a point-source level resolution collected at two LNG liquefaction sites, which we will refer to as Site A and Site B. The project started with a Baseline measurement campaign during which initial emissions measurements were taken. This was followed by several Enhanced Monitoring (EM) campaigns, each involving multiple

full facility scans. Each scan is a collection of individual Bridger passes (swaths) over the site, which together provide full coverage of the facility over a span of 1-2 hours. The project concluded with an End-of-Project (EOP) campaign at both sites. At Site A, the measurement campaigns spanned June 2022 to September 2023 and resulted in a total of 29 scans. At Site B, the measurement campaigns spanned October 2022 to September 2023 and resulted in a total of 20 scans. The exact dates and times of the measurement campaigns are listed in Section S1 of the Supporting Information (SI) document. Each measurement campaign (i.e., Baseline, EM, or EOP) typically comprised of one to six total facility scans conducted within a single day, except for the EOP campaigns, which spanned two days each. Note that Bridger used a rotary-wing aircraft during the BL and EOP campaigns, while a fixed-wing aircraft was employed during the EM campaigns. The different aircraft operate at varying altitudes and speeds, which influence the probability of detecting emissions.

We do not account for measurement uncertainties of the Bridger system, as the applicability of Bridger's measurement uncertainties has not been extensively evaluated in the context of LNG facilities, which differ significantly from production or midstream oil and gas sites. It is not clear if the methodologies proposed by Conrad et al.[34] for handling Bridger measurement uncertainties for measured emissions and by Johnson et al.[36] for unmeasured emissions are transferable to an LNG setting, particularly for measurements involving flares that in some cases extend over 500 feet tall. Therefore, we do not consider Bridger measurement uncertainty, and we also do not account for certain key emission sources, such as turbine exhaust, that Bridger does not detect.

Sites A and B have similar "functional elements," which are a collection of equipment groups involved in specific processes that perform analogous operational functions (e.g., LNG Tanks). Each functional element has multiple potential emission point sources. Our analysis focuses on emissions occurring directly on the sites, excluding those observed at offsite locations that were also captured by Bridger.

Note that Bridger reported methane emission rate estimates in standard cubic feet per

hour (scfh), which we convert to kg/hr for this paper. See details on unit conversion in SI Section S2. For confidentiality purposes, the data from both sites have been standardized using the corresponding site's data mean and standard deviation, that is we subtract the mean from all site measurements and divide by the standard deviation. This was done to anonymize the magnitude of the methane emissions while preserving the distributional behavior.

2.2 Data pre-analysis schemes

When a methane plume is detected on the edge of Bridger's measurement swath, Bridger is often able to identify elevated concentrations, indicating an emission, but not produce an emission rate estimate. Other technical or environmental conditions (e.g., rapidly changing wind direction) may also preclude an emission rate estimate when elevated concentrations are detected. The end-user of these data has several options for handling observed enhancements without a rate estimate. We consider the following non-exhaustive options: 1) exclude observations from the analysis when the Bridger algorithm lacks sufficient information to estimate an emission rate, 2) impute Bridger's minimum detection limit, which represents the smallest methane emission rate detectable with 90% certainty, under the assumption that the rate's absence is due to its low magnitude, and 3) impute the average emission rate observed at the specific source location during the corresponding measurement campaign. This third approach, similar to that used by Johnson et al.[36] for unmeasured emission sources, averages previous measurements instead of sampling from them. This results in the same mean but a narrower distribution compared to a sampling approach. The choice of method for handling these observations will affect the result of any subsequent analysis.

Furthermore, once an emission source is identified, Bridger will continue to monitor it in subsequent scans within the same measurement campaign. If a previously detected source does not show enhanced concentrations in later scans, Bridger will report "no gas detected." This outcome may indicate the intermittency of the emission, possibly because the source

ceased leaking by the time of the follow-up scan, or it could be due to the complexity of the LNG facility. For instance, a plume from one emission source could obscure another source positioned at a lower level of a large, multi-tier structure like an LNG Train.

The way we handle these non-detections, and the assumptions we choose to adopt, can significantly impact the resulting emission rate data. This decision-making process is referred to as the "emission persistence assumption." We consider two options:

1. Excluding "no gas detected" observations. This approach assumes that each source emits consistently, maintaining 100% persistence. Non-detections are considered anomalies, possibly due to detection limitations or environmental factors (e.g., complexity of wind patterns), rather than actual terminations of emissions and therefore are excluded from the analysis.

2. Imputing an emission rate of zero for "no gas detected" observations. Assigning an emission rate of zero to these cases suggests intermittent emissions, adjusting each source's persistence based on the frequency of these zero detections. A higher rate of non-detections indicates lower persistence, reflecting the sporadic nature of emissions.

We evaluate three data pre-analysis schemes, summarized in Table 1. A detailed description of each is as follows.

- *Data pre-analysis Scheme 1: Persistence Adjusted with 90% POD Substitution.* In this scheme, the adjusted persistence assumption imputes zero kg/hr rate estimates for "no gas detected" observations, which are then included in the calculation of point-source averages. This approach is equivalent to adjusting the Bridger average of positive emission rates by applying a persistence factor, defined as the ratio of the number of emission detections to the total number of passes within a scan. Additionally, when Bridger detects enhanced methane concentrations but fails to produce a rate estimate at the measurement swath (pass) level we impute minimum 90% probability of detection (POD) emission rate estimate, using 26 SCFH when Bridger's aircraft operates at

approximately 500 feet and 66 SCFH at altitudes above 700 feet.[35] Note that the POD also depends on the operating speeds of the aircraft.

- *Data pre-analysis Scheme 2: 100% Persistence without Substitution.* In this scheme, the 100% persistence assumption implies that emission source emits consistently, and thus "no gas detected" observations are considered as anomalies and excluded from the analysis. The observations where Bridger detects enhanced methane concentrations but fails to generate a rate estimate are also omitted and no substitution is applied.

- *Data pre-analysis Scheme 3: 100% Persistence with Average Substitution.* As in Scheme 2, the 100% persistence assumption implies that the emission source emits consistently. Therefore, any "no gas detected" observations are considered anomalies and are not included in the analysis. In situations where Bridger detects enhanced concentrations from a source but cannot quantify the emission rate, the average estimated emission rate from the corresponding campaign for that point source is used. If no such estimates are available, the emission rate is defaulted to the 90% POD rate.

Table 1: Summary of Bridger data pre-analysis schemes.

Data pre-analysis Scheme	**Substitution method for detected but not quantified sources**	**Persistence assumption**
1	90% POD	adjusted
2	excluded	100%
3	campaign average or 90% POD if no measurements are available	100%

Given that these non-exhaustive data pre-analysis schemes will affect the subsequent analysis, we first assess the impact of these three schemes on the resulting dataset. To see if the resulting datasets are statistically different, we use a non-parametric statistical test, the Kruskal-Wallis test, followed by post-hoc pairwise comparisons to test for significant differences between the resulting datasets. For these comparisons we use the pairwise Wilcoxon rank sum test, with a p-value adjustment applied to account for multiple comparisons.

2.3 Site-level aggregation methods

Currently in the US there are no regulatory specifications for calculating measurement-informed, site-level emissions inventories at oil and gas sites, including LNG sites. Average-based emission computation (e.g., temporal averages for functional elements) is intuitive for annual inventory reporting purposes. However, many top-down measurement technologies (e.g., satellite or aerial) provide a snapshot of a site's emissions, capturing data instantaneously without incorporating historical averages into a given measurement. With this in mind, our study compares two distinct approaches for aggregating methane emissions: 1) creating a distribution of temporal site-level total emission averages (such that represent annual site-level emissions) using an average-based method and 2) creating a distribution of instantaneous site-level total emission rates possible to observe at any given moment in time.

2.3.1 Average-based method for functional elements

One way that the aerial-measurements component of MII could be calculated is by using the average-based aggregation method. At a high-level, the method first calculates temporal averages for each point source (i.e., average point source emission rates within a scan), then creates spatial aggregation at the functional element-level (i.e., sum the point source averages within a given functional element), and finally takes a temporal average at the functional element level (i.e., average functional element-level estimates over multiple scans). The functional elements at LNG sites are often large equipment groups that contain multiple potential emission point sources, and as such, Bridger often detected multiple sources within each functional element.

In this method, we calculate the average emission rate for each point source during a scan by averaging all detected emissions from that specific point source across the duration of the scan. This results in one average emission rate per point source per scan. For each scan, we then aggregate the point source-level averages to the functional element-level. This

is done by summing the averaged point source-level emissions within each functional element for each scan. If a given functional element had no detected emissions during a given scan, we impute a 0 kg/hr emission for that scan. This results in one emission rate per functional element per scan. See section S3 in the SI for details on the imputed 0 kg/hr emissions.

To generate a distribution of potential site-level emission rates, we assume that emissions across all scans and equipment groups are independent. This simplification is necessary due to the limited availability of data, with only 29 scans at Site A and 20 scans at Site B. Assuming independence requires less data since we are not modeling the correlations or interactions between observations. Nevertheless, we evaluated an approach where emissions from different measurement campaigns, conducted at least a month apart, are treated as independent, while scan-level emissions within each campaign, typically lasting a day or two, are not. This method produced functional element-level distributions with similar means but drastically different shapes and wider spreads, attributable to small scan sample sizes (ranging from one to six scans within a given measurement campaign), see Section S4 in SI for details. However, this approach requires more aerial measurements to produce reliable site-level measurements. The operator noted that LNG facilities typically operate at a similar configuration (greater than 90% capacity factor), implying minimal operational differences among scans, regardless of whether they are conducted within the same or different measurement campaigns. Therefore, we opted to maintain the assumption of independence between scans for the analysis presented in this paper.

The average-based emission distributions that are based on the independence assumption of all scans are constructed for two LNG sites as follows:

1. For each functional element, randomly sample from the scan-level emission rates with replacement. Compute the average of the measurements within this sample. Repeat this process 10,000 times. For this method, we take the size of each sample to be the number of scans conducted at that site. This results in 10,000 bootstrapped average emission rates (across all scans) for each functional element. For example, randomly

draw with replacement 29 scan-level functional element emission rates (since Bridger conducted 29 scans of Site A) for Site A Train 1, compute their average emission rate, save this average value, and repeat 10,000 times. Do this for each Site A functional element. Follow the same steps for the Site B except using a sample size of 20 since Bridger conducted 20 scans of Site B. As per Central Limit Theorem, with a large enough sample size the distributions of the mean values will be normal unless the underlying distribution is highly skewed or sampling process is not truly random.

2. For each functional element, fit a normal distribution to the 10,000 bootstrapped overall average emission rates from the previous step.

3. For each functional element, take 10,000 samples from the normal distributions that were fitted in the previous step. This approach has several benefits: (1) it models the underlying distribution of a functional element, extending beyond the limited scope of the observed samples; (2) sampling from this fitted distribution smooths out anomalies due to random variation or outliers, thus yielding a more generalized view of the distribution; (3) it enables the simulation of data points that, while not originally present in the dataset, are statistically consistent with the existing data.

4. Sum the samples across functional elements, working element-wise. This results in 10,000 bootstrapped, site-level emission rate estimates.

Most of the bootstrapped functional element-level average emission rates were well approximated by a normal distribution. However, the distributions for East Jetty and Train 9 were not well described by the Normal distribution, as the underlying distributions were multi-modal. The distributions for Marine Flare, BOG Compressors, West Jetty, and LNG Tanks 2 were not perfectly symmetric. We tested a non-linear kernel fit as an alternative to the normal fits (see Section S5 in SI for details), but the differences were minimal, leading us to conclude that the Normal distribution adequately describes the distribution of overall means for each functional element. This outcome is expected with an appropriate sample

size and when the underlying distributions are not highly skewed. If normality cannot be assumed for the sample mean distribution, steps 2 and 3 mentioned above must be modified.

2.3.2 Instantaneous method

In this method, we estimate a distribution of instantaneous emission rates. This distribution describes the relative probability of observing a given site-level emission rate during any snapshot measurement of the full site. The average-based method aggregates emissions at the functional element level and assumes that all average emission rates are independent and equally likely. The instantaneous method instead samples directly at the individual point source level and assumes that all point source-level emission rates are independent and equally likely.

To account for intermittency at the point source level, we impute zero emission rates for the scans in which Bridger did not detect a given source. For example, during the EM2 measurement campaign, the Site A Marine Flare underwent four scans, yet no emissions were detected. Consequently, we impute four instances of zero kg/h emission rates for the Marine Flare point source, reflecting the absence of emissions during those scans. As a result, at Site A, every point source has a minimum of 29 emission rates, with more in instances where the point source was observed multiple times within a given scan. During a scan each point source can be observed multiple times due to overlaps between the measurement swaths (passes).

To produce the distribution of instantaneous rates, we first randomly sample one emission rate, which may include rates of zero kg/hr, from each detected point source at a site. Next, we sum these rates across all point sources within the site to calculate a site-level emission rate. This process is repeated 10,000 times, generating 10,000 estimates of the site-level instantaneous emission rates.

This procedure imposes the following assumptions, which we consider a reasonable starting point. Future work will relax some of these assumptions.

1. The presence and magnitude of an emission at one location are not correlated with the presence and magnitude of an emission at any other location.

2. There is no overlap between the Bridger measurement swaths. In practice, certain methane emission point sources may be measured more than once during a single scan of the entire facility due to overlap between the individual swaths. Therefore, in our current analysis, we are likely including fewer zero emission rates than we should be, which biases the distribution upwards and results in overall higher estimates.

3 Results

The following subsections will demonstrate the differences in data pre-analysis schemes and in site-level aggregation methods. Specifically, we will highlight how average-based methods do not capture the variability of instantaneous emissions.

3.1 Data pre-analysis schemes

Figure 1 shows the point source-level average emission rates (i.e., no aggregation method applied) resulting from the three different pre-analysis schemes discussed in Section 2.2. The most conservative data pre-analysis scheme, "Data pre-analysis Scheme 3", results in the highest site-level emission rates, as expected. This is because 100% persistence is assumed, and either average campaign emission rates or 90% Bridger POD were imputed for cases where gas is detected but no rate is produced. The pairwise Wilcoxon rank sum test reveals that there is a statistically significant difference between Schemes 1 and 2 and between Schemes 1 and 3 (p-value $<$ 2e-16). No statistical difference was found between Scheme 2 and Scheme 3 at the 0.05 significance level (p-value = 0.14). The median of the data from Scheme 1 is lower than those of the other two pre-analysis schemes. This is because Scheme 1 adjusts for persistence, while Schemes 2 and 3 assume 100% persistence. We conclude, therefore, that the persistence assumption is the most important pre-analysis consideration.

Differences between Schemes 2 and 3 are minor, meaning that the method for handling the no-quantification detections has minimal impact on the resulting emission rate distribution.

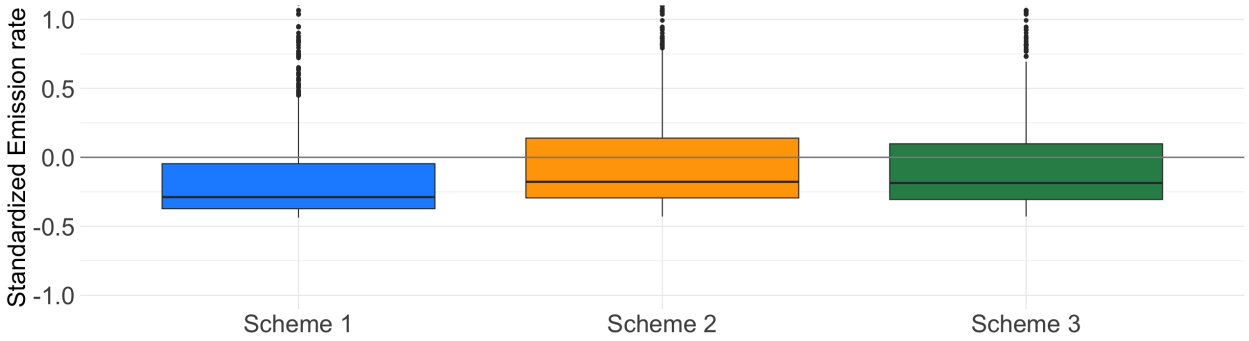


Figure 1: Point source-level average emission rates (i.e., no aggregation method applied) resulting from the three considered data pre-analysis schemes discussed in Section 2.2. Data are shown as boxplots, with whiskers showing 1.5 times the interquartile range. The y-axis is limited to [-1, 1] to show detail.

Note that we do not recommend a specific data pre-analysis scheme; rather, we highlight that the choice of pre-analysis scheme can significantly influence the resulting emissions datasets.

3.2 Site-level aggregation methods

3.2.1 Average-based Method for Functional Elements

For the remainder of this section, we show results using "Data pre-analysis Scheme 3," as it retains the highest amount of information among the available options.. The two sites considered in this analysis have similar functional elements and operational processes. However, the major contributors to overall emissions differed between the two sites. We use the average-based method to evaluate the contributions of each functional element to the overall site-level emissions.

Figure 2 shows each functional element's scan-level average emission rate as a percent of scan-level total site emissions. The average emissions are computed prior to imputing scan-level zero emission rates. For Site A Flares are the primary source of methane emissions

in nearly every scan. Additionally, there is large variability in flare emissions between the different scans. Because the flare emissions account for the largest portion of scan-level site emissions, the variability in flare emissions also drives the site-level emissions variability between scans. The second largest source of emissions at Site A is the Marine Flare. There was a major operational change that occurred at the Marine Flare between Scans 7 and 8. According to information from the operator, before this date the marine flare at Site A had been utilized during every cooldown and, briefly, during regular ship loading. Since the operational changes were made, the marine flare has been used during ship loading only when the vapor from the ship is off-spec (i.e., the composition has amount of CO2 and water vapor beyond the site's tolerance), resulting in emissions that are an order of magnitude lower than before this operational change.

At Site B, on the other hand, Trains have the highest methane emissions during all scans. Similar to the flares at Site A, the variability in emission rates observed between scans at Trains is the primary factor influencing the fluctuations in the estimated site-level emission rates at Site B. The Flares and Marine Flares on Site B have lower magnitude compared to site A and also contribute less to scan-level site emissions, due to the presence of large emissions from the Trains.

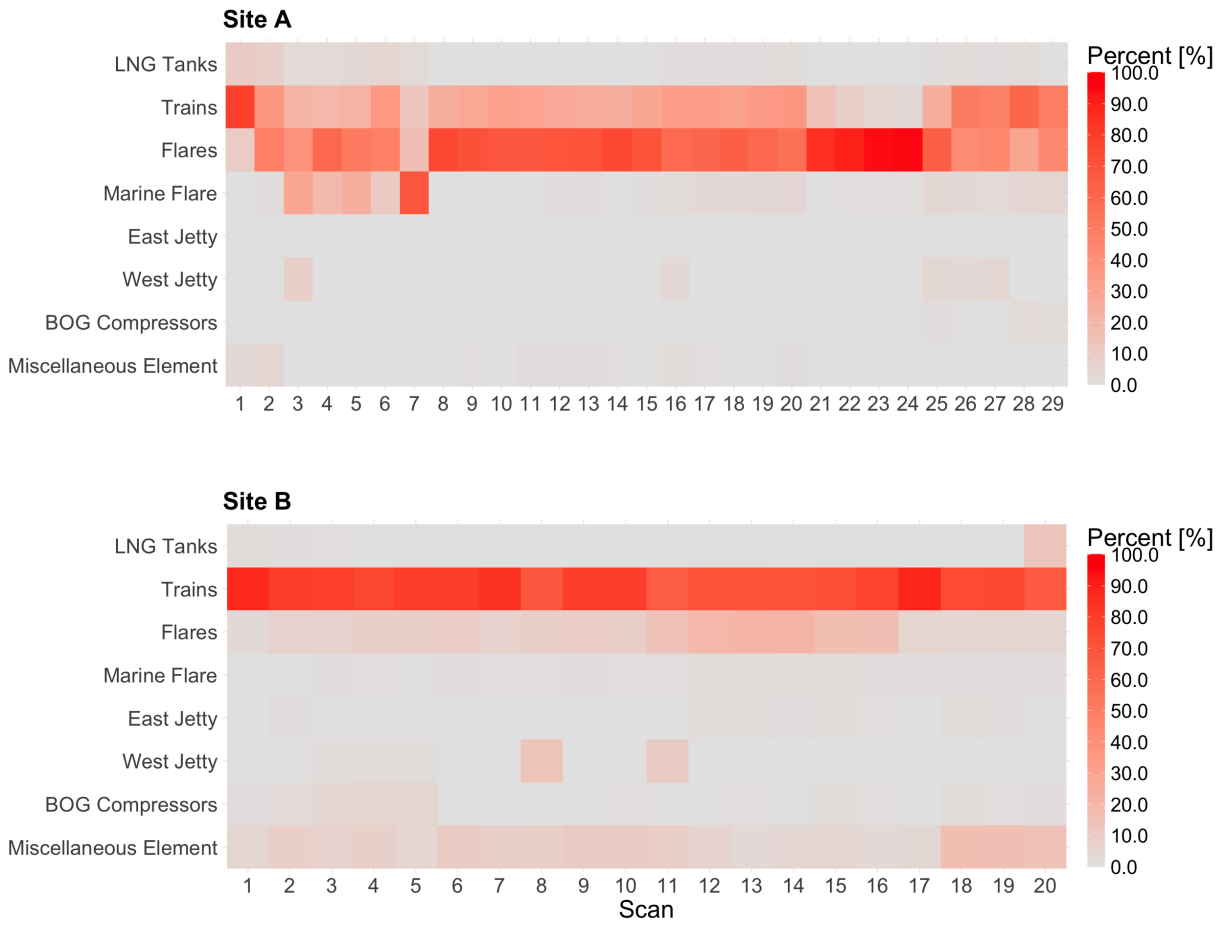


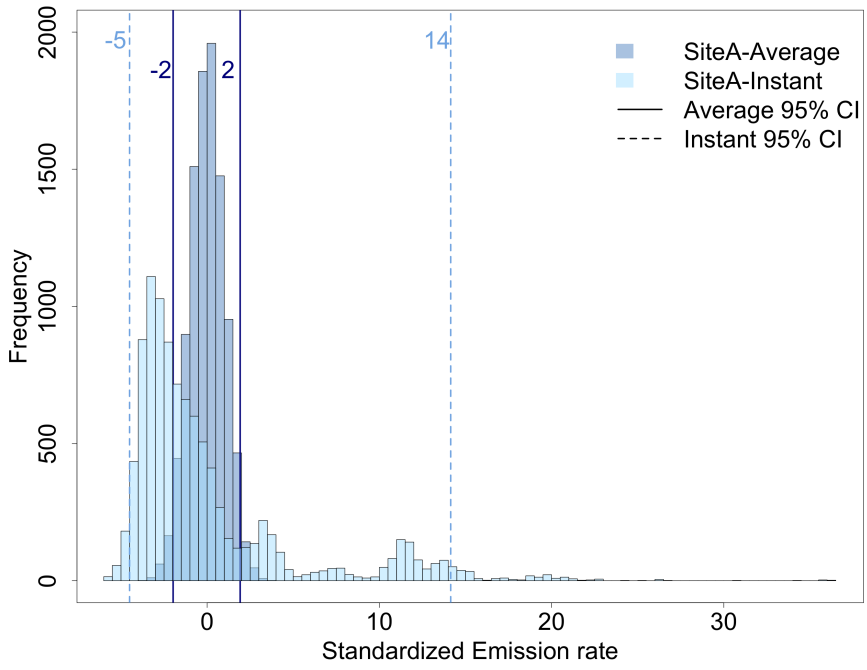
Figure 2: The percent contribution of each functional element's average emission rate to the scan-level total emissions at each facility prior to scan-level zero emission rates being imputed.

3.2.2 Comparative analysis of instantaneous and average-based methods

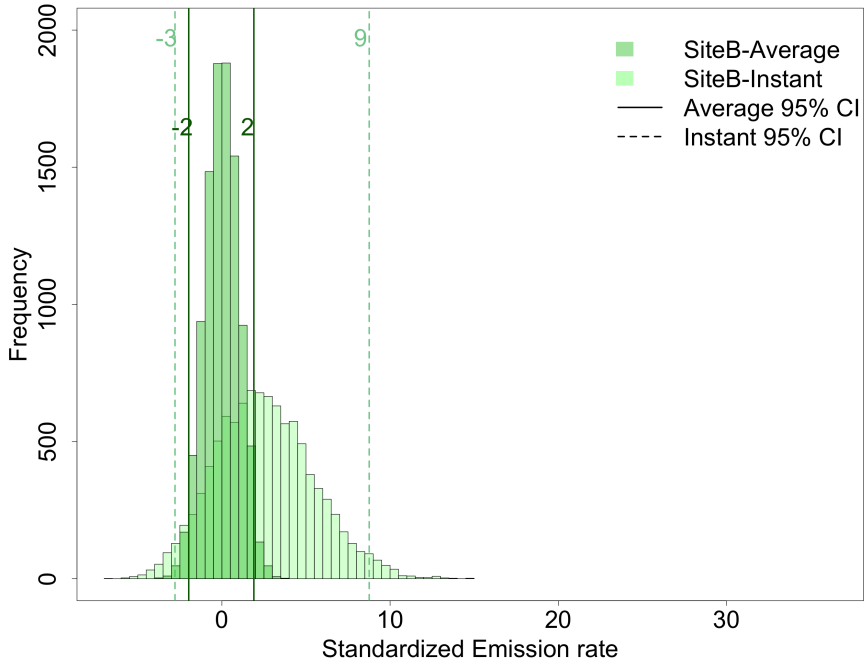
Figure 3 shows the distribution of site-level emission rates at Sites A and B, derived from both the instantaneous and average-based approaches. Recall that the data from both sites have been standardized using the corresponding site average-based distribution parameters, thus preserving the emission distribution shapes of both sites while ensuring comparability between average-based and instantaneous distributions. The shape and spread of these distributions give one measure of uncertainty in the site-level rate estimates. The standardized average emissions at Sites A and B are normally distributed, with both having a 95% confi-

dence interval of [-2, 2] as expected. The instantaneous distributions of both sites have heavy tails, which is driven by the presence of infrequent large emissions in the underlying point source-level data. The 95% confidence interval, [-5, 14], of Site A instantaneous site-level emissions is wider than the 95% confidence interval, [-3, 9], of Site B instantaneous site-level emissions. The instantaneous method draws from emissions point sources directly and hence more often includes large, sporadic emissions that are smoothed out when the averaging aggregation method is used.

As expected, we see that the instantaneous distributions are wider than the average distributions on both sites, as averages dampen the influence of infrequent large emission detections. This implies that a snapshot measurement from, for example, a satellite or an aerial survey may show emissions that are considerably larger or smaller than described by an average-based inventory. Specifically, the standard deviation of the emissions inferred from the average-based method is 5 times smaller than the standard deviation of the emissions inferred from the instantaneous method on site A, and 3 times smaller on site B.



(a) Site A



(b) Site B

Figure 3: Site-level emissions distributions using both the averaging and instantaneous methods for Sites A and B. The 95% confidence intervals are identified using solid lines for average site-level emissions and dashed lines for instantaneous site-level emissions. Emission rates have been standardized for anonymity.

4 Discussion

In this study, we aim to bridge a gap in the literature by providing a comprehensive analysis of site-level methane emissions at LNG facilities. We use two aggregation methods to calculate emissions at the site level using data from the QMRV project collected on two LNG facilities. Key learnings from this project include the following:

1. Data pre-analysis significantly influences the emission estimates. Specifically, adjusting for persistence compared to assuming 100% persistence has a statistically significant impact (p-value $<$ 2e-16) on resulting emission estimates. Excluding detected but not quantified emissions compared to using the 90% POD or campaign averages did not result in a significant difference (p-value = 0.14).

2. Similar LNG facilities operated by the same company can have different emission characteristics. Although the standardized emission rates for Sites A and B were similar in shape and magnitude when calculated using the average-based method, the standardized emissions at Site A have a much wider spread than those at Site B when assessed with the instantaneous method. Furthermore, the flares were the largest contributors to emissions on Site A, while trains were the largest contributors on Site B.

3. There is substantial variation in emissions between functional elements on a given site. Therefore, measurements at the functional element-level, followed with further root cause analysis at the source level, are necessary for targeted emission mitigation strategies.

4. The emission rate estimates derived from multiple scans highlight the intermittent nature of emissions from key components, particularly flares and trains. Properly accounting for this intermittency or understanding its causes is crucial for developing accurate, measurement-based emission inventories at the site level.

5. We develop a method for estimating the distribution of instantaneous site-level emissions. We find that the distribution of instantaneous emissions is much wider than the distribution of average emissions for both sites. This implies that snapshot measurements, which capture the instantaneous emission rate of an LNG facility, can fall outside of the distribution describing the annual inventory (i.e., the annual mean) and the uncertainty surrounding it. The instantaneous approach should not be used for MII computations, because it describes the variability of the emissions at specific moments in time, rather than providing an accurate representation of the annual inventory.

This study highlights the variability in methane emissions at LNG facilities and the complexity of aggregating them at the site-level. Leveraging data from the LNG QMRV project, we have demonstrated the impact of data pre-analysis schemes on emission estimates. Additionally, we create and further investigate two distinct methodologies for calculating site-level emissions: the average-based method, which represents the variability in the average site-level emission rate, and the instantaneous method, which represents the variability in the possible emission rates on the site at any given point in time. We would like to emphasize that the instantaneous approach should not be used for inventory purposes, such as annual emissions reporting, whether for regulatory compliance or voluntary disclosures, as it doesn't represent annual site-level emissions. We find that these methods yield different distributions of site-level emissions, with the average-based method smoothing the instantaneous emission variability. This analysis, predicated on the assumption that emission events are independent, simplifies the complexity of emission dynamics and may not fully reflect the actual behavior of emissions at a site. Future research will quantify the correlation structure of methane emissions on LNG facilities to relax this assumption. Additionally, we will extend this analysis to other types of oil and gas facilities, such as those in the midstream and production sectors.

Funding

This work was funded by Cheniere Energy, Inc.

Competing Interests

F.C.G., and G.B.R. are employees of Cheniere Energy Inc. S.A.R.W was an employee of Cheniere till May 2024, and is currently an employee of SLR International. SLR International performs work for Cheniere, other oil and gas industry clients, academic institutions, and industry research organizations. A.P.R. is currently a member of the Gas Pipeline Advisory Committee of the US Department of Transportation; in this role, he is a Special Government Employee. A.P.R. has current research support from the US Department of Energy, Environmental Defense Fund, and sponsors of the Energy Emissions Modeling and Data Lab (EEMDL). D.M.H. has current research support from the US Department of Energy, NASA, and EEMDL sponsors.

Acknowledgement

The authors thank the participating LNG facility site operators, and the measurement technology vendor, Bridger, for their contributions to this research project.

Supporting Information Available

The Supporting Information document includes the following sections:

- Measurement campaigns.
- Bridger measurement unit.
- Imputation of zero emission rates.

- Independence of measurement campaigns.
- Normal and kernel fit.

References

(1) Safari, A.; Das, N.; Langhelle, O.; Roy, J.; Assadi, M. Natural gas: A transition fuel for sustainable energy system transformation? *Energy Science & Engineering* **2019**, *7*, 1075–1094, https://doi.org/10.1002/ese3.380.

(2) (EIA), E. I. A. Natural Gas Explained: Natural Gas and the Environment. U.S. Energy Information Administration. EIA, 2022; https://www.eia.gov/energyexplained/natural-gas/natural-gas-and-the-environment.php, https://www.eia.gov/energyexplained/natural-gas/natural-gas-and-the-environment.php.

(3) Allen, D. T.; Torres, V. M.; Thomas, J.; Sullivan, D. W.; Harrison, M.; Hendler, A.; Herndon, S. C.; Kolb, C. E.; Fraser, M. P.; Hill, A. D.; Lamb, B. K.; Miskimins, J.; Sawyer, R. F.; Seinfeld, J. H. Measurements of methane emissions at natural gas production sites in the United States. *Proceedings of the National Academy of Sciences* **2013**, *110*, 17768–17773, https://doi.org/10.1073/pnas.1304880110.

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

(5) Zimmerle, D. J.; Williams, L. L.; Vaughn, T. L.; Quinn, C.; Subramanian, R.; Duggan, G. P.; Willson, B.; Opsomer, J. D.; Marchese, A. J.; Martinez, D. M.; Robinson, A. L. Methane Emissions from the Natural Gas Transmission and Storage System in the United States. *Environmental Science & Technology* **2015**, *49*, 9374–9383, https://doi.org/10.1021/acs.est.5b01669.

(6) Alvarez, R. A. et al. Assessment of methane emissions from the U.S. oil and gas supply chain. *Science* **2018**, *361*, 186–188, https://doi.org/10.1126/science.aar7204.

(7) Wang, J. L.; Daniels, W. S.; Hammerling, D. M.; Harrison, M.; Burmaster, K.; George, F. C.; Ravikumar, A. P. Multiscale Methane Measurements at Oil and Gas Facilities Reveal Necessary Frameworks for Improved Emissions Accounting. *Environmental Science & Technology* **2022**, *56*, 14743–14752, https://doi.org/10.1021/acs.est.2c06211.

(8) Rutherford, J. S.; Sherwin, E. D.; Ravikumar, A. P.; Heath, G. A.; Englander, J.; Cooley, D.; Lyon, D.; Omara, M.; Langfitt, Q.; Brandt, A. R. Closing the methane gap in US oil and natural gas production emissions inventories. *Nature Communications* **2021**, *12*, 4715, https://doi.org/10.1038/s41467-021-25017-4.

(9) Sherwin, E. D.; Rutherford, J. S.; Zhang, Z.; Chen, Y.; Wetherley, E. B.; Yakovlev, P. V.; Berman, E. S. F.; Jones, B. B.; Cusworth, D. H.; Thorpe, A. K.; Ayasse, A. K.; Duren, R. M.; Brandt, A. R. US oil and gas system emissions from nearly one million aerial site measurements. *Nature (London)* **2024**, *627*, 328–334D, https://doi.org/10.1038/s41586-024-07117-5.

(10) Intergovernmental Panel On Climate Change (IPCC) *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, 1st ed.; Cambridge University Press, 2023; https://doi.org/10.1017/9781009157896.

(11) Ocko, I. B.; Sun, T.; Shindell, D.; Oppenheimer, M.; Hristov, A. N.; Pacala, S. W.; Mauzerall, D. L.; Xu, Y.; Hamburg, S. P. Acting rapidly to deploy readily available methane mitigation measures by sector can immediately slow global warming. *Environmental Research Letters* **2021**, *16*, 054042, https://doi.org/10.1088/1748-9326/abf9c8.

(12) United Nations Environment Programme and Climate and Clean Air Coalition *Global Methane Assessment: Benefits and Costs of Mitigating Methane Emissions*; United Nations Environment Programme: Nairobi, 2021; https://www.unep.org/resources/report/global-methane-assessment-benefits-and-costs-mitigating-methane-emissions.

(13) European Parliament Reducing Methane Emissions in the Energy Sector. https://www.europarl.europa.eu/legislative-train/theme-a-european-green-deal/file-reducing-methane-emissions-in-the-energy-sector, 2024; Accessed: 2024-05-03.

(14) Williams, C.; DiSavino, S. US was top LNG exporter in 2023 as hit record levels. Reuters, 2024; https://www.reuters.com/business/energy/us-was-top-lng-exporter-2023-hit-record-levels-2024-01-02/.

(15) Innocenti, F.; Robinson, R.; Gardiner, T.; Howes, N.; Yarrow, N. Comparative Assessment of Methane Emissions from Onshore LNG Facilities Measured Using Differential Absorption Lidar. *Environmental Science & Technology* **2023**, *57*, 3301–3310, https://doi.org/10.1021/acs.est.2c05446.

(16) Omara, M.; Zavala-Araiza, D.; Lyon, D. R.; Hmiel, B.; Roberts, K. A.; Hamburg, S. P. Methane emissions from US low production oil and natural gas well sites. *Nature Communications* **2022**, *13*, 2085, https://doi.org/10.1038/s41467-022-29709-3.

(17) Subramanian, R.; Williams, L. L.; Vaughn, T. L.; Zimmerle, D.; Roscioli, J. R.; Herndon, S. C.; Yacovitch, T. I.; Floerchinger, C.; Tkacik, D. S.; Mitchell, A. L.; Sullivan, M. R.; Dallmann, T. R.; Robinson, A. L. 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* **2015**, *49*, 3252–3261, https://doi.org/10.1021/es5060258.

(18) United Nations Environment Programme Oil and Gas Methane Partnership. https://ogmpartnership.com/, 2024; Accessed: 2024-05-03.

(19) Handler, B.; Ayaburi, F. The cleaning of U.S. natural gas; evolution of differentiated gas and related crediting mechanisms. *Resources Policy* **2024**, *90*, 104750, https://doi.org/10.1016/j.resourpol.2024.104750.

(20) Erland, B. M.; Thorpe, A. K.; Gamon, J. A. Recent Advances Toward Transparent Methane Emissions Monitoring: A Review. *Environmental Science & Technology* **2022**, *56*, 16567–16581, https://doi.org/10.1021/acs.est.2c02136.

(21) 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* **2014**, *119*, 6836–6852, https://doi.org/10.1002/2013JD021272.

(22) Johnson, M. R.; Tyner, D. R.; Conley, S.; Schwietzke, S.; Zavala-Araiza, D. Comparisons of Airborne Measurements and Inventory Estimates of Methane Emissions in the Alberta Upstream Oil and Gas Sector. *Environmental Science & Technology* **2017**, *51*, 13008–13017, https://doi.org/10.1021/acs.est.7b03525.

(23) Irakulis-Loitxate, I. et al. Satellite-based survey of extreme methane emissions in the Permian basin. *Science Advances* **2021**, *7*, eabf4507, https://doi.org/10.1126/sciadv.abf4507.

(24) Feitz, A. et al. The Ginninderra CH4 and CO2 release experiment: An evaluation of gas detection and quantification techniques. *International Journal of Greenhouse Gas Control* **2018**, *70*, 202–224, https://doi.org/10.1016/j.ijggc.2017.11.018.

(25) El Abbadi, S. H.; Chen, Z.; Burdeau, P. M.; Rutherford, J. S.; Chen, Y.; Zhang, Z.; Sherwin, E. D.; Brandt, A. R. Technological Maturity of Aircraft-Based Methane

Sensing for Greenhouse Gas Mitigation. *Environmental Science & Technology* **2024**, acs.est.4c02439, https://doi.org/10.1021/acs.est.4c02439.

(26) Dooley, J. F.; Minschwaner, K.; Dubey, M. K.; El Abbadi, S. H.; Sherwin, E. D.; Meyer, A. G.; Follansbee, E.; Lee, J. E. A New Technique for Airborne Measurements to Quantify Methane Emissions Over a Wide Range: Implementation and Validation. 2024; https://doi.org/10.5194/egusphere-2024-760.

(27) Daniels, W. S.; Wang, J. L.; Ravikumar, A. P.; Harrison, M.; Roman-White, S. A.; George, F. C.; Hammerling, D. M. Toward Multiscale Measurement-Informed Methane Inventories: Reconciling Bottom-Up Site-Level Inventories with Top-Down Measurements Using Continuous Monitoring Systems. *Environmental Science & Technology* **2023**, *57*, 11823–11833, https://doi.org/10.1021/acs.est.3c01121.

(28) Chen, Q.; Schissel, C.; Kimura, Y.; McGaughey, G.; McDonald-Buller, E.; Allen, D. T. Assessing Detection Efficiencies for Continuous Methane Emission Monitoring Systems at Oil and Gas Production Sites. *Environmental Science & Technology* **2023**, *57*, 1788–1796, https://doi.org/10.1021/acs.est.2c06990.

(29) Daniels, W. S.; Jia, M.; Hammerling, D. M. Detection, localization, and quantification of single-source methane emissions on oil and gas production sites using point-in-space continuous monitoring systems. *Elem Sci Anth* **2024**, *12*, 00110, https://doi.org/10.1525/elementa.2023.00110.

(30) Bridger Photonics, I. Gas Mapping LiDAR for Aerial Methane Detection. 2024; https://www.bridgerphotonics.com/gas-mapping-lidar-for-aerial-methane-detection, Accessed: 2024-05-16.

(31) Johnson, M. R.; Tyner, D. R.; Szekeres, A. J. Blinded evaluation of airborne methane source detection using Bridger Photonics LiDAR. *Remote Sensing of Environment* **2021**, *259*, 112418, https://doi.org/10.1016/j.rse.2021.112418.

(32) Bell, C.; Rutherford, J.; Brandt, A.; Sherwin, E.; Vaughn, T.; Zimmerle, D. Single-blind determination of methane detection limits and quantification accuracy using aircraft-based LiDAR. *Elementa: Science of the Anthropocene* **2022**, *10*, 00080, https://doi.org/10.1525/elementa.2022.00080.

(33) Rutherford, J.; Sherwin, E.; Chen, Y.; Aminfard, S.; Brandt, A. Evaluating methane emission quantification performance and uncertainty of aerial technologies via high-volume single-blind controlled releases. 2023; https://doi.org/10.31223/X5KQ0X.

(34) Conrad, B. M.; Tyner, D. R.; Johnson, M. R. Robust probabilities of detection and quantification uncertainty for aerial methane detection: Examples for three airborne technologies. *Remote Sensing of Environment* **2023**, *288*, 113499, https://doi.org/10.1016/j.rse.2023.113499.

(35) Thorpe, M. J. et al. Deployment-invariant probability of detection characterization for aerial LiDAR methane detection. **2024**, http://dx.doi.org/10.2139/ssrn.4762406.

(36) Johnson, M. R.; Conrad, B. M.; Tyner, D. R. Creating measurement-based oil and gas sector methane inventories using source-resolved aerial surveys. *Communications Earth & Environment* **2023**, *4*, 139, https://doi.org/10.1038/s43247-023-00769-7.

(37) Schuller, O.; Kupferschmid, S.; Hengstler, J.; Whitehouse, S. Life cycle GHG emission study on the use of LNG as marine fuel. *Hinkstep: Stuttgart, Germany* **2019**, https://sea-lng.org/wp-content/uploads/2020/06/19-04-10_ts-SEA-LNG-and-SGMF-GHG-Analysis-of-LNG_Full_Report_v1.0.pdf.

(38) Okamura, T.; Furukawa, M.; Ishitani, H. Future forecast for life-cycle greenhouse gas emissions of LNG and city gas 13A. *Applied Energy* **2007**, *84*, 1136–1149, https://doi.org/10.1016/j.apenergy.2007.05.005.

(39) Tamura, I.; Tanaka, T.; Kagajo, T.; Kuwabara, S.; Yoshioka, T.; Nagata, T.; Kura-

hashi, K.; Ishitani, H. Life cycle CO2 analysis of LNG and city gas. *Applied Energy* **2001**, https://doi.org/10.1016/S0306-2619(00)00062-3.

(40) Zhang, Z.; Cusworth, D. H.; Ayasse, A. K.; Sherwin, E. D.; Brandt, A. R. Measuring Carbon Dioxide Emissions From Liquefied Natural Gas (LNG) Terminals With Imaging Spectroscopy. *Geophysical Research Letters* **2023**, *50*, e2023GL105755, https://doi.org/10.1029/2023GL105755.

(41) Zhu, Y.; Ross, G.; Khaliukova, O.; Roman-White, S. A.; George, F. C.; Hammerling, D. M.; Ravikumar, A. P. Multi-scale Measurements of Greenhouse Gas Emissions at U.S. Natural Gas Liquefaction Terminals. *ChemRxiv* **2024**, https://doi.org/10.26434/chemrxiv-2024-h4flq.

(42) Schwietzke, S. et al. Improved Mechanistic Understanding of Natural Gas Methane Emissions from Spatially Resolved Aircraft Measurements. *Environmental Science & Technology* **2017**, *51*, 7286–7294, https://doi.org/10.1021/acs.est.7b01810.

(43) Cheniere Energy, I. A Closer Look at QMRV. 2024; https://www.cheniere.com/our-responsibility/climate/closer-look-qmrv, Accessed: 2024-05-16.

TOC Graphic