**Supplemental Information**

Characterizing detection probabilities of advanced mobile leak surveys: implications for sampling effort and leak size estimation in natural gas distribution systems

Emily Luetschwager1, Joseph C. von Fischer2, Zachary Weller1\*

1 Department of Statistics, Colorado State University, Fort Collins CO USA

2 Department of Biology, Colorado State University, Fort Collins CO USA

\*zach.weller@colostate.edu

List of Contents:

**Text S1. Empirical variation in excess CH4 measurements.**

**Figure S1.** **Boxplots showing the observed distribution of CH4 enhancements.**

These boxplots show the distribution of log CH4 enhancements for 42 CH4 sources that were each detected on at least 40 different drive-bys. The CH4 sources are sorted by their estimated average emission rate. For most sources, there is a long upper tail of CH4 enhancements, even after the log transformation. This long tail can lead to overestimation as detailed in the main text.

**Figure S2.** **Boxplots demonstrating the range of within-source CH4 enhancement variation**.

The boxplots show the range of variation for 12 sources, each with at least 10 detections. For each of the empirical min, 33rd percentile, 66th percentile, and max of within-leak CH4 enhancement standard deviations, we chose three representative CH4 sources.

**Text S2. Further results of the logistic regression analysis.**

**Table S1.** **Parameter estimates and profile confidence intervals for the detection probability analysis**

**Text S3. Sensitivity analysis for modeling detection probabilities.**

**Figure S3.** **Sensitivity analysis of detection probability analysis.**

The inclusion of singleton detections decreases detection probability across all cities and leak sizes. In some cities, the inclusion of singletons causes the estimated detection probability to decrease with increasing emission rate.

**Text S4. Additional results from the Monte Carlo synthesis.**

**Table S2**. **Counts of observed peaks (OP) and verified peaks (VP) from our simulation as a function of sampling effort**

**References**

*S1. Empirical Variation in Excess CH4 Measurements.*

The results of our percent difference in size estimation simulation demonstrated that CH4 sources can produce expressions that make them appear as large or small leaks. Here, we provide further results documenting the variation in CH4 enhancements that occur from in-field CH4 sources. These results demonstrate both the variability of within-source CH4 enhancements and the variation of within-source variability across sources seen in our surveys.

Figure S1 displays boxplots of the log CH4 enhancements for 42 CH4 sources. Each of these sources were detected on 40 or more drive-bys during our surveying. The sources are ranked by their estimated emission rates and colored by the standard deviation of the log CH4 enhancements. Hereafter, log represents log10 unless specified otherwise. A notable feature of many of these sources is that, even on the log scale, the distribution of excess CH4 concentrations still has a right skew. As noted in the main text, this can cause positive bias in estimated emission rates.



**Figure S1.** **Boxplots showing the observed distribution of CH4 enhancements.** These boxplots show the distribution of log CH4 enhancements for 42 CH4 sources that were each detected on at least 40 different drive-bys. The CH4 sources are sorted by their estimated average emission rate. For most sources, there is a long upper tail of CH4 enhancements, even after the log transformation. This long tail can lead to overestimation as detailed in the main text.

Figure S2 displays the boxplots of the log CH4 enhancements for 12 leaks that were chosen to demonstrate the range of within-leak variability seen in our surveys. Each of these sources were detected on at least ten drive-bys and they are ranked by their standard deviation. Sources 1-3 correspond to the three largest standard deviations, and sources 10-12 correspond to the three smallest standard deviations. Source 5 corresponds to the 66th percentile of within-leak standard deviations, and sources 4 and 6 are the sources immediately above and below (respectively) the 66th percentile. Similarly, source 8 corresponds to the 33rd percentile of within-leak standard deviations, and sources 7 and 9 are the sources immediately above and below (respectively) the 33rd percentile.

Figure S2 illustrates the range and variability of within-leak standard deviations. We see that the within leak standard deviation varies by more than 3x. This variation is likely due to differences in source emission rate, proximity to the survey vehicle, and other environmental factors mentioned in the main text.

About 25% (2164/8191) of the CH4 sources detected in our survey had a larger standard deviation than the standard deviation observed in the Picarro instrument calibration experiment (Von Fischer et al., 2017). In our calibration experiment, our estimated within-leak standard deviation for log CH4 enhancements was 0.39. These large standard deviations suggest that further controlled release experiments are needed to capture the range of variation seen in the field.



**Figure S2.** **Boxplots demonstrating the range of within-source CH4 enhancement variation**. The boxplots show the range of variation for 12 sources, each with at least 10 detections. For each of the empirical min, 33rd percentile, 66th percentile, and max of within-leak CH4 enhancement standard deviations, we chose three representative CH4 sources.

*S2. Further Results of the Logistic Regression Analysis.*

Here we provide further details on the detection probability analysis and results. We estimated the following logistic regression model:

where, denotes city/urban area , enumerates verified CH4 sources in city , is the estimated source emission rate (log L/min), is a fixed intercept across all the cities, is a fixed slope across all cities, and is a random, city-specific intercept. We assume that the city-specific intercepts follow a normal distribution, . We estimated model parameters using maximum likelihood. Estimated model parameters and their confidence intervals are given in Table S1.

There are some limitations of our logistic regression analysis. First, our model assumes that the estimated emission rates are known. Second, our analysis also does not account for environmental factors that affect detection. In some cases, these factors are unknown (e.g., distance between the vehicle and the expression location) and in other cases we did not have the capability to measure these factors (e.g., wind). Finally, we only considered verified CH4 sources (sources seen on at least two drive-bys) for our analysis. We provide a sensitivity analysis of this choice in the next section.

**Table S1.** **Parameter estimates and profile confidence intervals for the detection probability analysis**

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Estimate** | **95% CI** |
| Fixed Intercept () | -0.566 | (-0.77,-0.36) |
| Fixed Slope (Size Estimate) () | 0.821 | (0.77,0.87) |
| Random Effect Intercept SD (City Variation) | 0.338 | (0.24,0.54) |

The slope parameter and corresponding confidence intervals suggest that larger leaks are more likely to be detected. We provide discussion about leak detection in the main text. There is also evidence of variability between city-specific intercepts, as indicated by the confidence interval for the standard deviation of these intercepts. This variation is displayed in Figure 3 in the main text.

*S3. Sensitivity analysis for modeling detection probabilities.*

In our detection probability analysis, we included only CH4 sources that had been verified, meaning they had been detected at least twice at a given location. We use this criterion in our reporting in order to reduce the number of false positive leak indications. For the purposes of detecting NG pipeline leaks, we consider a false positive as any elevated CH4 readings that arise due to instrument error or CH4 sources that are non-attributable to a pipeline leak (e.g., CNG buses, biological CH4 sources). An open question regarding ALD surveys is the rate of false positives.

In practice, some NG pipeline leaks will only be detected once and never verified. We can also observe elevated CH4 concentrations at a location only once because of the presence of CH4 sources that are not pipeline leaks. We describe CH4 sources that have been detected only once as singletons. Some unknown fraction of these singletons are false positives and the remaining correspond to NG pipeline leaks.

We conducted a sensitivity analysis to examine the effects of including singletons in our detection probability analysis. Some results of this sensitivity analysis are shown in Figure S3. In Figure S3, the first panel displays the results of fitting logistic regression models to each city separately using only verified peaks. The middle panel displays the same results but where we have included singletons up to three drives (e.g., any source that was detected at least once on three or fewer drive-bys).  Similarly, the right panel displays the results where we have included singletons up to ten drives. As expected, these results demonstrate that estimated detection probabilities decrease when singletons are included. In some cities, the inclusion of singletons causes estimated detection probabilities to decrease with increasing leak size. These results suggest that our estimated detection probabilities are positively biased due to the exclusion of singletons, but the degree to which they may be biased is unclear due to the lack of understanding of our false positive rate, wind effects, and leak expression locations.



**Figure S3.** **Sensitivity analysis of detection probability analysis.** The inclusion of singleton detections decreases detection probability across all cities and leak sizes. In some cities, the inclusion of singletons causes the estimated detection probability to decrease with increasing emission rate.

*S4. Additional results from the Monte Carlo synthesis.*

From the Monte Carlo synthesis, we documented the number of leaks that only are observed once and not verified (singletons). As indicated in Table 1 in the main text, there is a large proportion of singleton OPs from our surveying in 15 cities. In contrast to these results, in our simulation we saw a much smaller proportion of singleton OPs across all levels of sampling effort. This occurs because we did not consider false positive detections in our simulation and is further evidence that we may be biased in our estimation detection probabilities by not including singletons in our analysis. The results again highlight the need to evaluate the false-positive rate/frequency of non-NG sources that are detected in the mobile surveys.

**Table S2**. **Counts of observed peaks (OP) and verified peaks (VP) from our simulation as a function of sampling effort**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Drives | OP Count | VP Count | Proportion (count)  of non-verified OPsa | Overall VP detection probability |
| 2 | 172 | 83 | 0.52 (89) | 0.73 |
| 3 | 189 | 139 | 0.27 (50) | 0.75 |
| 4 | 195 | 169 | 0.13 (26) | 0.71 |
| 5 | 198 | 185 | 0.07 (13) | 0.68 |
| 6 | 199 | 192 | 0.03 (7) | 0.67 |
| 7 | 200 | 196 | 0.02 (3) | 0.65 |
| 8 | 200 | 198 | 0.01 (2) | 0.66 |
| 9 | 200 | 199 | 0 (1) | 0.67 |
| 10-15 | 200 | 200 | 0 (0) | 0.66 +/- 0.01 |

aComputed proportion of observed peaks that were not part of a verified peak in our simulation.

We note that the proportion of non-verified OPs is much lower than the empirical proportions from our survey results. The overall VP detection probability from our simulation is higher than the detection probabilities from our survey results.

*References*

Von Fischer JC, Cooley D, Chamberlain S, Gaylord A, Griebenow CJ, Hamburg SP, Salo J, Schumacher R, Theobald D, Ham J. 2017. Rapid, Vehicle-Based Identification of Location and Magnitude of Urban Natural Gas Pipeline Leaks. *Environ Sci Technol* **51**(7): 4091–4099. American Chemical Society. doi: 10.1021/acs.est.6b06095