Supporting Information

Tiered Leak Detection and Repair Programs at Simulated Oil and Gas Production Facilities: Increasing Emission Reduction by Targeting High-Emitting Sources

Felipe J. Cardoso-Saldaña*

ExxonMobil Technology and Engineering Company, Spring TX 77389, USA

*** Corresponding Author:**

Felipe J. Cardoso-Saldaña

ExxonMobil Technology and Engineering Company

22777 Springwoods Village Parkway, Spring, TX, 77389, USA

E-mail: felipe.j.saldana@exxonmobil.com

Number of pages: 25

Number of figures: 16

Number of tables: 2

Table of Contents

S.1. Facilities in Simulation

A total of 96 tank batteries taken from Stokes et al.[1] were used in the model, including their equipment counts. In addition to tank batteries, wellhead only sites were also included and thus it was necessary to estimate their number. The first step was to estimate the number of wells included in wellhead only sites. The total number of tanks at the tank batteries was added and divided by a ratio of 0.84 tanks per well from Ravikumar[2] to estimate the total number of wells. The number of wells that need to be assigned to wellhead only sites was estimated by subtracting the number of wells at tank batteries from the total number of wells. The second step was to estimate the number of wellhead only sites. Rutherford et al.[3] developed scaling factors for equipment per well for the natural gas and the petroleum systems using data from the US Environmental Protection Agency (EPA) Greenhouse Gas Reporting Program (GHGRP) and reported 0.84 meters/piping per gas well and 0.22 headers/piping per oil well. Because the number varies per type of well, here the percentage of oil and gas wells in the Permian basin was estimated based on 2020 GHGRP data[4] as 95% and 5%, respectively. The number of sites was estimated by assuming one piping at each wellhead only site: oil sites were estimated by multiplying the total number of wells to add by 0.95 and by 0.22, whereas for gas wells the total number of wells to add was multiplied by 0.5 and by 0.84. This lead to a total of 161 wellhead only sites and a total of 257 facilities simulated.

The model tracks whether components are emitting or not throughout the simulation, and thus it was necessary to estimate the number of components at each equipment and site. Rutherford et al.[3] reports a range of number of components per equipment type for each type of equipment present in the natural gas system and in the petroleum system from a compilation of multiple sources. Here, the middle point from that range was taken and multiplied by the number of equipment at each site to estimate the number of components. For equipment present both in oil and gas systems, a weighted average of the middle point values in the range was estimated using 0.95 for the oil system and 0.5 for the gas system. For piping at wellhead only sites, the lowest number in the range from Rutherford et al.[3] was used.

S.2 Emission Measurements

In the main paper it was described that the percentage of sites with leaks from the close-range inspection datasets was 71%. The total number of sites surveyed by Bridger Photonics (1251) was

multiplied by 0.71 to estimate the facilities expected to have had emissions below the aerial detection threshold, leading to 888. The number of leaks present at these sites was estimated by randomly sampling, with replacement, the number of leaks per site for sites that had detections in the close-range studies. The site approach was preferred than a leak per well approach because some tank batteries do not have co-located wells. Emission rates from the various close-range studies were aggregated into one list of measurements, and the total number of leaks at the 888 sites was sampled from the measurements with replacement, to get the emission rates below the aerial surveys detection threshold. This random sampling was performed at each Monte Carlo iteration in the simulation, generating a slightly different set of measurements from which to sample for emission rates in each Monte Carlo run.

Some of the largest emissions from the close-range field studies were on the range of what Bridger Photonics can observe in the field, and thus to avoid double counting these emissions from those that were observed in the aerial surveys, emissions were assigned a probability of being removed based on the detection curve in Eq. S1 to account for emissions within the partial detection range of Bridger Photonics. This approach is similar to Chen et al.[5]

$$Probability\ of\ detection = 0.5 * (1 + \text{erf}\left(\left(\frac{(leak\ rate - m}{\sigma}\right)/\sqrt{2}\right)) \quad \text{(Eq. S1)}$$

Where erf is the error function, leak rate is the leak rate in units of scfh, m is 106.5 and σ has a value of 26.9.

S.3. Leak Generation Rates

S.3.1. Deriving Emission Duration Based on Cusworth et al.

Cusworth et al.[6] performed extensive aerial surveys of sites in the Permian Basin in 2019 using imaging spectrometers. Sites were visited multiple times over a period of six weeks. Since each facility was visited an average of 8 times and the survey period was short, it was possible to use these data to estimate the duration of emissions above the detection threshold of the measurements (10-20 kg/hr). Cusworth et al.[6] report dates when emission were observed. The dates when emissions were not detected were not reported, but can be inferred from the reported data.

To estimate the numbers of non-detects by site, polygons were constructed for each sampling day using the coordinates of the facilities that reported emissions with the 'st_convex_hull' function from the 'sf' package in the R programming language. These polygons were then perturbed to match non-overlapping daily flight regions that align with the areas that Cusworth et al.[6] described as the flight areas for the GAO aircraft for both the Delaware and Midland basins. Figure S1 shows a schematic of this process for the Delaware basin. Sites whose coordinates lay within the areas flown, for which no emissions were reported, were assumed to have no detected emissions on that day. It is possible that some facilities would have been surveyed more than once on a given day if there was an overlap on adjacent flight lines. However, given that the shapefiles containing the flight polygons by Cusworth et al.[6] are not publicly available, it is not possible to know which facilities might have been sampled more than once in a given day. Limiting the analysis to daily flyovers introduces bias, and the duration of emissions might be less than estimated here.

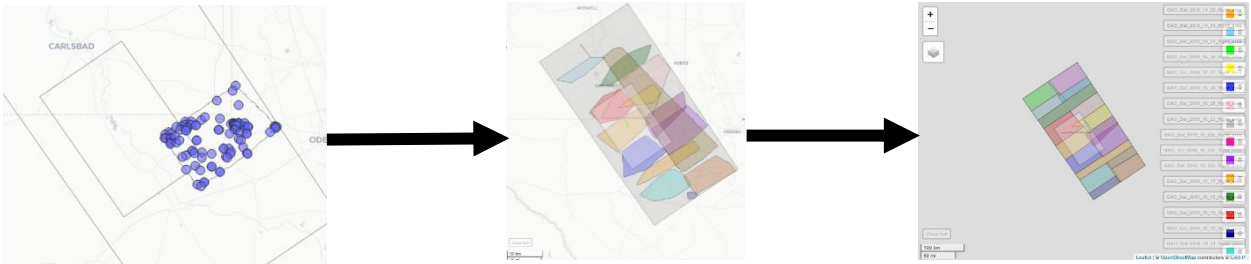


Figure S1. Identification of areas flown on each day by Cusworth et al.[6]

The accuracy of this methodology was assessed by comparing Cusworth, et al.[6]'s reported persistence to a persistence calculated based on the inferred non-detections. To estimate persistence, for each site, the number of days with emission detection were added to the days with no detected emissions, to arrive at an estimated total number of flights. The number of detections divided by the estimated total number of flights is the estimated persistence. The persistence reported by Cusworth, et al.[6] was compared to the estimated persistence for sites with more than 3 flyovers. Comparisons are shown in Figure S2 ($R^2 = 0.83$).

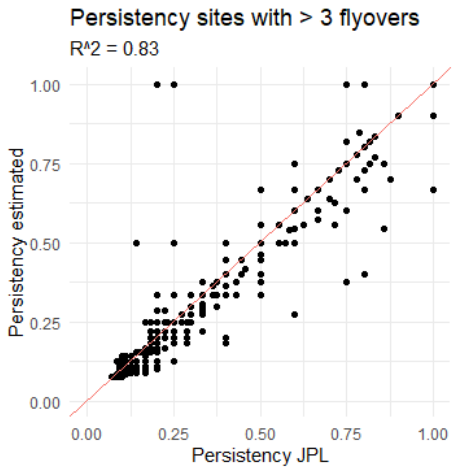


Figure S2. Estimated vs reported persistence for sites with more than 3 flyovers. The 1:1 line is shown in red.

Time series of emissions were constructed based on the dates when emissions were either detected or not detected at each facility. Days between flights were assumed to be emitting if sequential flights had detected emissions. Days between flights were assumed to not be emitting if sequential flights did not have detected emissions. When, for sequential flights, one flight detected emissions and one did not, the emissions were assumed to persist for half of the time between flights. Days before the first flight and after the last flight for each site were not considered.

These time series were used to create a distribution of estimated emission durations (Figure S3). A large fraction of emissions have durations <1-2 days, peaks are observed at around 5 days and around 10 days, and there are relatively few sites with emissions lasting more than 12 days. Distributions can also be created by facility type. Pipelines and compressor stations tend to have lower proportions of emissions lasting <1-2 days, while well sites and tank batteries have the highest proportion of durations <1-2 days.

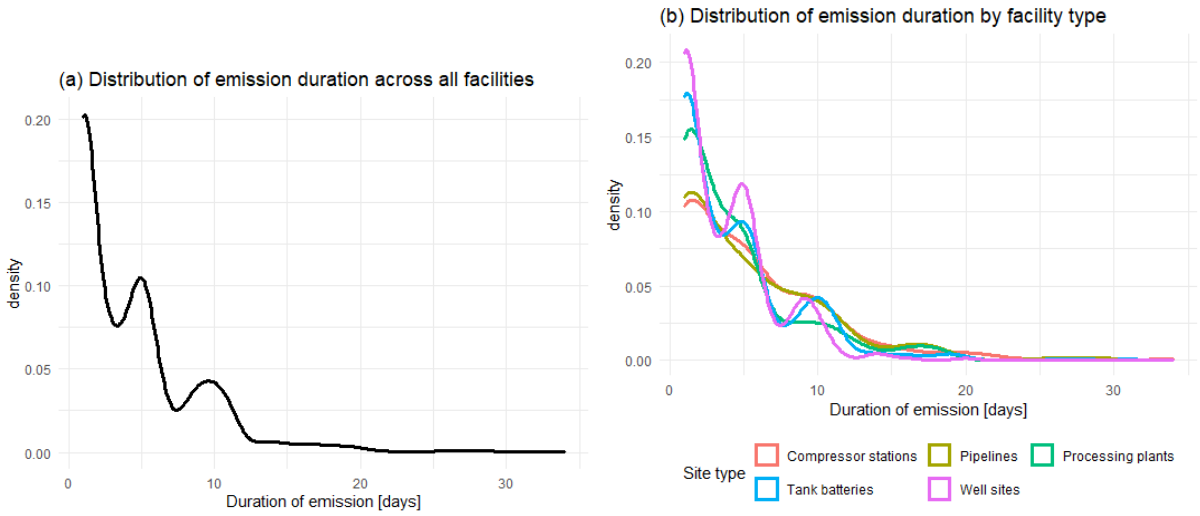


Figure S3. Time duration of emissions derived from Cusworth et al.[6] for (a) all facilities and (b) segregated by facility type.

There are multiple uncertainties associated with this analysis, including, but not limited to the identification of dates with non-detects, overlapping flight plumes leading to facilities having more than one survey in a given day, the lack of data on the duration of emission detections occurring on either the first or last measurement at a site, and biases introduced by unequal times between sequential flights. Nevertheless, mean and median values of persistence provide some guidance regarding durations of events. Median values of emission durations are estimated to be 3-5 days (Table S1), with wells and tank battery sites having shorter durations, and compressor stations and pipelines having longer durations. To test the robustness of these results to the assumptions made in the analyses, the number of days with flights with no detection was increased or decreased, selecting a day at random, such that the estimated persistence agreed exactly with the persistence reported by Cusworth et al.[6] This procedure was repeated 100 times and the mean and median for each updated distribution with their 95% confidence interval are shown in Table S1. The mean of all facilities is 4.9 days, with a range of 4 to 5.8 days. A sensitivity analysis S3 was performed by assigning average duration of 5 days (MTTR) to emissions from tanks and flares (Table 2), which are the two largest sources of high-emitters.[7]

Table S1. Mean and median emission duration by type of facility derived from Cusworth et al.[6]

Type of facility	Mean	Median
All	4.9 (4.8-5)	4 (3–4)
Tanks	4.8 (4.6–5)	3 (3–4)
Pipelines	5.6 (5.3–5.9)	4 (4–5)
Processing	4.9 (4.3–5.6)	3 (3–4)
Compressor	5.8 (5.7–6.1)	5 (4–5)
Well	4 (3.8–4.1)	3(3–3)

The field study from Cusworth et al.[6] observed both routine and un-intended emissions. Thus, the estimate average duration from observations is a combination of short-duration routine emissions (e.g. compressor start-ups and blow-downs), which drive the average persistency down; of un-intended emissions that were not fixed during the field campaign if operators didn't have an LDAR program at the time of the campaign; and of un-intended emissions that could have been repaired in the field if operators were having either LDAR or audible visual olfactory (AVO) surveys and repairs during that time. To account for the fact that routine emissions could be driving the average down, and to have more data points varying leak generation rate values, two additional sensitivity analyses were performed on the simulations: assuming an MTTR for tanks and flares of 10 days (S4) and of 30 days (S5) as shown on Table 2. Leak generation rates are a difference in the modeling here done compared to other open source LDAR models; in those models the same leak generation rate is used independently on the component or equipment type.[8,9]

S.3.2 Distribution of Times to Leak and Stop Leaking

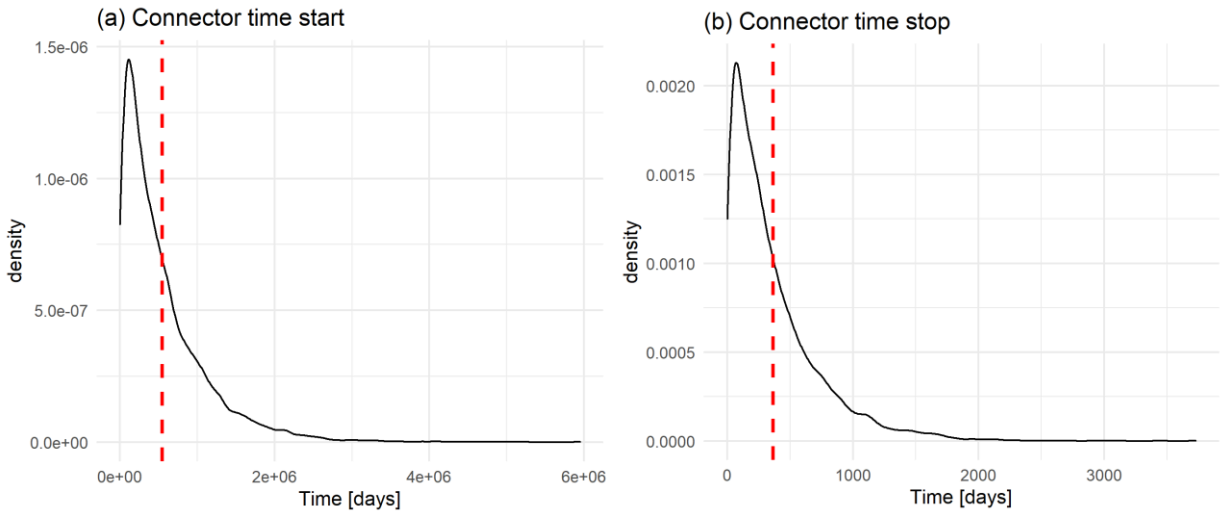


Figure S4. Distribution of times to start of next leak (a) and time for leak to stop outside LDAR inspections (b) for connectors. The MTBF and MTTR values for connectors from Table 1 are shown as dashed red lines.

S.3.3. Emission Duration Effect on Reduction

Figure S5 shows emission time series where two hypothetical groups of sites have the same number of emitters at each time step (pLeak same on both). In this example, group 1 has longer duration on emissions and group 2 a shorter duration. The group of sites with longer emission duration leads to higher reduction, whereas the group with shorter emission duration leads to lower reduction since emissions are appearing and disappearing quickly across various sites. This schematic shows that to have reductions on emissions, the timeframe of LDAR inspections needs to be less than the timeframe of the emission duration.

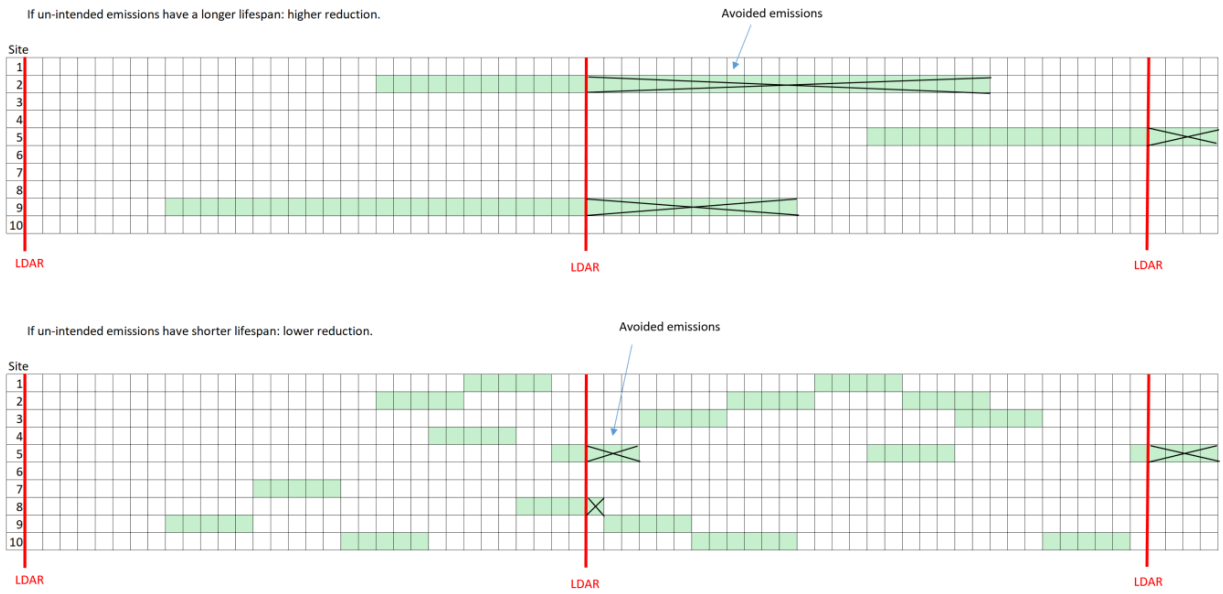


Figure S5. Time series of emissions at hypothetical groups of facilities that have the same number of emitters at any given point in time, but with different durations of emissions.

S.4. Cloud Coverage

Mean number of clear, partly cloudy and cloudy days for Midland, TX, were retrieved from the US National Centers for Environmental Information.[10] A clear day indicates a fraction of the sky covered between 0 and 3 (out of 10), a partly cloudy one represents a fraction of the sky covered between 4 and 7 (out of 10), while a cloudy day has a fraction of the sky covered between 8 and 10 (out of 10). As the simulation progresses over time, monthly data from the day under simulation is retrieved, and a cloud coverage is randomly sampled, first whether clear, partly cloudy and cloudy, and then the fraction of the sky covered. At each day in the simulation, a random number between 1 and 10 is picked; if the number is larger than the fraction of the sky covered in that day, it was assumed that the satellite will be able to detect emissions that day, otherwise the satellite detections are not implemented that day.

S.5. Simulations Ran

A total of 566 different scenarios of LDAR programs were ran. Table S2 shows a condensed summary of the simulations ran. For each row of Table S2, all combinations that can be created based on the various frequencies and detection thresholds were simulated.

Table S2. Simulations included in scenarios.

Survey type	Frequency	Detection threshold
OGI	1,2,4,6 and 12x	Zimmerle et al.[11] high experience of operators.
Aerial + OGI	1-11x aerial + 1x OGI	Aerial: 2, 5, 10 and 25 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.
Satellite + aerial + OGI	Satellite daily and weekly revisit + 1-11x aerial + 1x OGI	Satellite: 50 and 100 kg/hr. Aerial: 2, 5, 10 and 25 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.
Satellite + OGI	Satellite daily and weekly revisit + 1x OGI	Satellite: 50 and 100 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.
Continuous network + OGI	1x OGI	Continuous: 5 and 10 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.
Satellite + continuous network + OGI 1x	Satellite daily and weekly revisit + 1x OGI	Satellite: 50 and 100 kg/hr. Continuous: 5 and 10 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.
Continuous site level + OGI	1x OGI	Continuous: 0.2, 2, 5, 10 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.
Satellite + continuous site level + OGI	Satellite daily and weekly revisit + 1x OGI	Satellite: 50 and 100 kg/hr. Continuous: 0.2, 2, 5, 10 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.
Continuous on priority sites + OGI	1x OGI	Continuous: 0.2, 2, 5, 10 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.
Continuous on priority sites + aerial on non-priority + OGI	1-5x aerial + 1x OGI	Continuous: 0.2, 2, 5, 10 kg/hr. Aerial: 2 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.
Satellite + continuous on priority sites + OGI	Satellite daily and weekly revisit + 1x OGI	Satellite: 50 and 100 kg/hr. Continuous: 0.2, 2, 5, 10 kg/hr.

		OGI: Zimmerle et al.[11] high experience of operators.
Satellite + continuous on priority sites + aerial on non-priority + OGI	Satellite daily and weekly revisit + 1-5x aerial + 1x OGI	Satellite: 50 and 100 kg/hr. Continuous: 0.2, 2, 5, 10 kg/hr. Aerial: 2 kg/hr. OGI: Zimmerle et al.[11] high experience of operators.

Each of the 566 simulations that are condensed in Table S2 were simulated in the base case and in 6 additional sensitivity analyses, for a total of 3962 model runs. Table 2 in the main text shows the summary of parameters varied in sensitivity analyses compared to the base case simulations.

S.6. Effect of Including Flares in the Simulation

Emissions from flares were modeled in a simplified way, and thus it was important to understand what impact might arise from including them in the results, in particular when comparing the performance of LDAR programs with respect to each other (on a relative basis). For each simulation ran, the emission reduction was estimated with and without inclusion of emissions from flares (in both the no-LDAR runs which were used as a baseline to estimate reduction and on the LDAR runs), and the results plotted in Figure S6a for the base case scenarios, and S6b for the S3 sensitivity analysis scenarios. In both plots, the data points lay very close to the 1:1 line suggesting that the impact of including emissions from flares vs not including them is not significant. The OGI 4x reduction lines are shown as black lines, and the data points lay within the top right and bottom left quadrants generated by the OGI 4x lines, suggesting that any LDAR simulation that has a larger reduction than OGI 4x when considering flares also has a larger reduction than OGI 4x when not considering flares, which leads to consistent results. There would be disagreements when comparing LDAR programs on a relative basis if some scenarios had reductions larger than OGI 4x with flares and lower than OGI 4x without flares, or vice versa. Since including flares does not affect results when comparing LDAR programs on a relative basis and it has small variation in the absolute emission reduction, in this work the comparisons were done including emissions from flares.

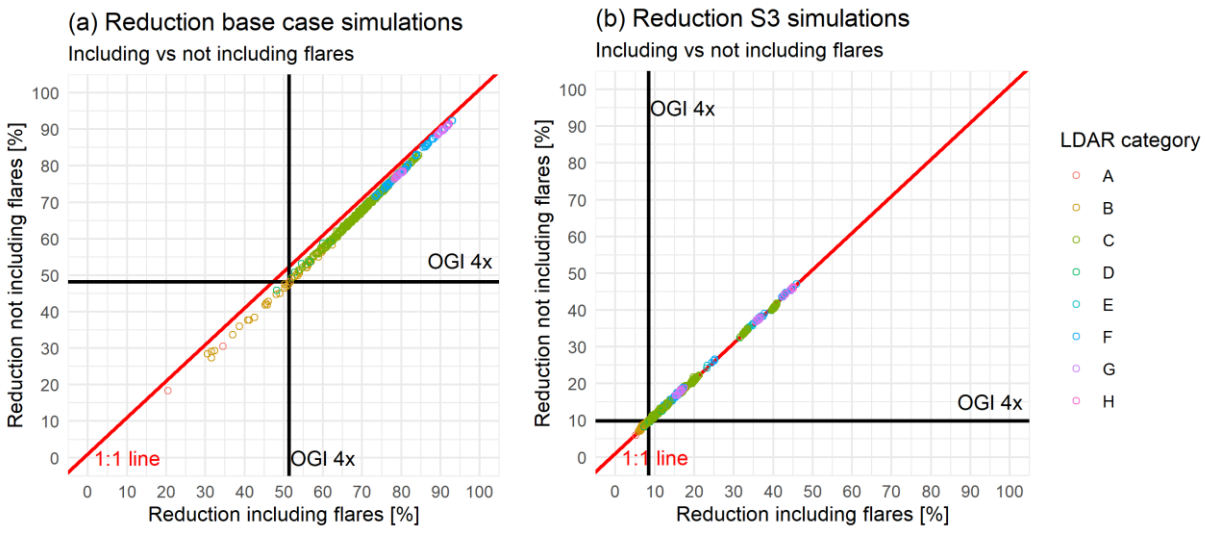


Figure S6. Reduction of simulations with and without flares for the (a) base case and (b) sensitivity analysis S3 simulations. The red line indicates the 1:1 line. The data points are colored by LDAR category: "A" = OGI, "B" = Aerial + OGI, "C"= Satellite + aerial + OGI, "D" = Satellite + OGI, "E" = Continuous + OGI, "F" = Satellite + continuous + OGI, "G" = Satellite + continuous at priority sites + OGI, "H" = Satellite + continuous at priority sites + aerial at non priority sites + OGI.

S.7. Assessing Sensitivity of Model Parameters in Emission Reduction

An important step in the analysis was to understand which model parameters had a larger effect on the emission reduction. Figure S7 shows the distribution of the difference in emission reduction between the 566 simulations in each sensitivity analysis with respect to those in the base case. Sensitivity analysis with a distribution near zero have a minor effect on the results, whereas those that range farther than zero have a larger impact. The results of the S0 sensitivity analysis scenarios are very similar with respect to the base case ones, the mean difference in the reduction percentage is -0.8 (+/- 0.6), indicating that the choice of studies to aggregate to create the emission distribution for small emitters does not have a large impact in addition to sampling independently of the equipment where the emission originated. S1 and S2 have a mean difference in reduction percentage of -1 (+/- 0.7) and -0.9 (+/-0.7) with respect to the base case results, suggesting that when combining the distributions of close-range inspections and flyover emissions, the fraction of sites with emissions from the close-range inspections (smaller emission rates) is not as sensitive.

On the other hand, S3, S4 and S5 sensitivity analyses differ more significantly from the base case, by -53.5 (+/- 8.3), -41.9 (+/- 9) and -20.1 (+/- 6.1), respectively, suggesting that the duration of high-emitters and the rate at which they appear are highly sensitive in the model. These results are in line with Fox et al.[9] The remaining of the analysis was focused on both the base case and S3, given that these two are the most extreme scenarios, to understand how reductions can vary based on the duration of high emitters.

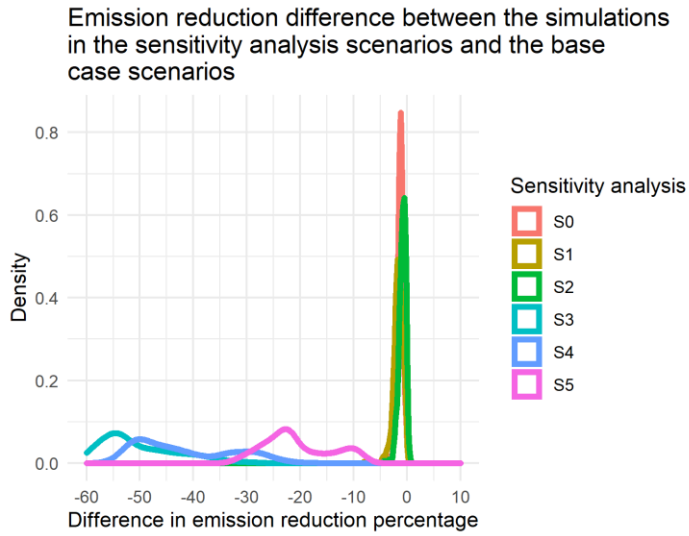


Figure S7. Emission reduction difference of the sensitivity analyses simulations with respect to those in the base case. S0 are the scenarios where the emission measurements were assigned independently of the equipment of the leak and additional close-range datasets were included. S1 and S2 varied the number of facilities with emissions from the close-range inspection studies when combining these measurements with the aerial flyover measurements. S3 was the shortest duration of high-emitters. S4 was a short duration of high emitters, but longer than S3. S5 was a duration of high-emitters shorter than the base case and longer than S4.

S.8. Combinations of Satellite, Continuous Monitoring, Aerial and OGI Sensors

S.8.1. Site Level Continuous Monitoring

For site level continuous monitors, when additional satellite-based detections are included, there is no significant improvement in emissions in the scenarios with long duration of emissions (Figure S8a). However, higher reductions are achieved when satellites are coupled with continuous monitoring sensors in the scenarios of low duration of high-emitters, and this difference is larger than the difference in reduction that is due to a decrease in the detection threshold of the sensors (Figure S8b). This is explained because in the simulation, emissions detected by a satellite were assumed to be repaired within 2 days, while those with continuous monitors within 7 days. All scenarios achieve higher reduction than the OGI only scenarios. While decreasing the detection threshold of the continuous monitors increases the emission reduction, all of them achieve more reduction than 12x OGI, and their difference is relatively small particularly in scenarios of short duration of high-emitters.

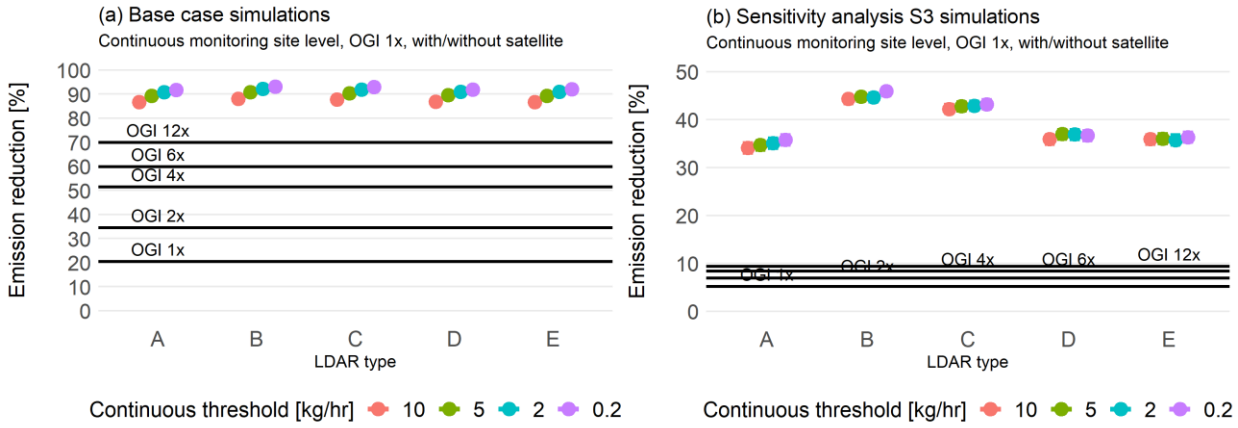


Figure S8. Emission reduction of LDAR scenarios with satellite + continuous monitoring site level + OGI for (a) base case (long duration of high-emitters) and (b) sensitivity analysis S3 (short duration of high-emitters) simulations. LDAR type A is no satellite. LDAR type B has a satellite with 50 kg/hr and daily revisits. LDAR type C has a satellite with 100 kg/hr and daily revisits. LDAR type D has a satellite with 50 kg/hr and weekly revisits. LDAR type E has a satellite with 100 kg/hr and weekly revisits. Horizontal black lines indicate the reduction of OGI only LDAR programs.

S.8.2. Site Level Continuous Monitoring at Priority Sites

In scenarios that had continuous monitoring at priority sites (tank batteries), the non-priority sites (wellhead only) were simulated under various frequencies of aerial based detection. The results show that there is no difference on reductions on having additional flyovers on these facilities, besides the yearly OGI, because the emission profile of these sources is very low (Figure S9).

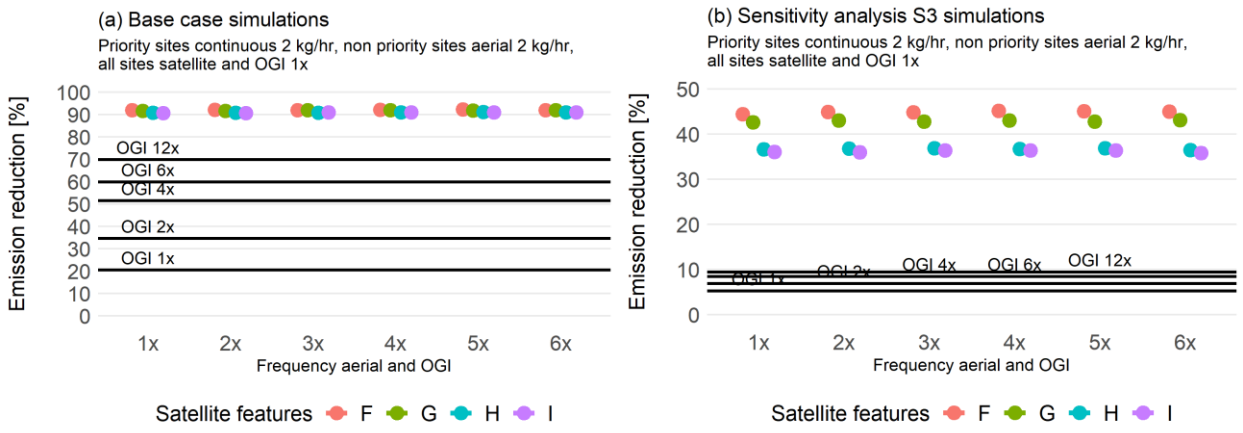


Figure S9. Emission reduction of LDAR scenarios with satellite + continuous monitoring site level on priority sites + aerial on non-priority sites + OGI for (a) base case (long duration of high-emitters) and (b) sensitivity analysis S3 (short duration of high-emitters) simulations. "F" is satellite with detection threshold of 50 kg/hr and daily revisit. "G" is satellite with detection threshold of 100 kg/hr and daily revisit. "H" is satellite with detection threshold of 50 kg/hr and weekly revisit. "I" is satellite with detection threshold of 100 kg/hr and weekly revisit. Horizontal black lines indicate the reduction of OGI only LDAR programs The horizontal axis represents the frequency of aerial and OGI, and OGI is used only once a year in all scenarios so the variation on frequency is due to the aerial surveys.

S.9. Effect of Repair Times

Figure S10 shows the emission reductions for aerial 10 kg/hr + satellites + OGI 1x. Figures S10c and S10d show the effect of repairing all leaks within 30 days, vs Figures S10a and S10b fixing those found with satellites within 1-2 days. Across all scenarios the reductions are lower than with emissions prioritized, although having a tiered approach with a satellite is better than only aerial + 1x OGI. In scenarios of long duration of high emitters, the quick repair of emissions detected with satellites can help achieve reductions above 12x OGI for many of the scenarios, while most of the scenarios with repair within 30 days still do better than 4x OGI but less than 12x OGI. The effect

is similar in scenarios of short duration of high-emitters, although all of them perform better than 12x OGI when repaired quickly.

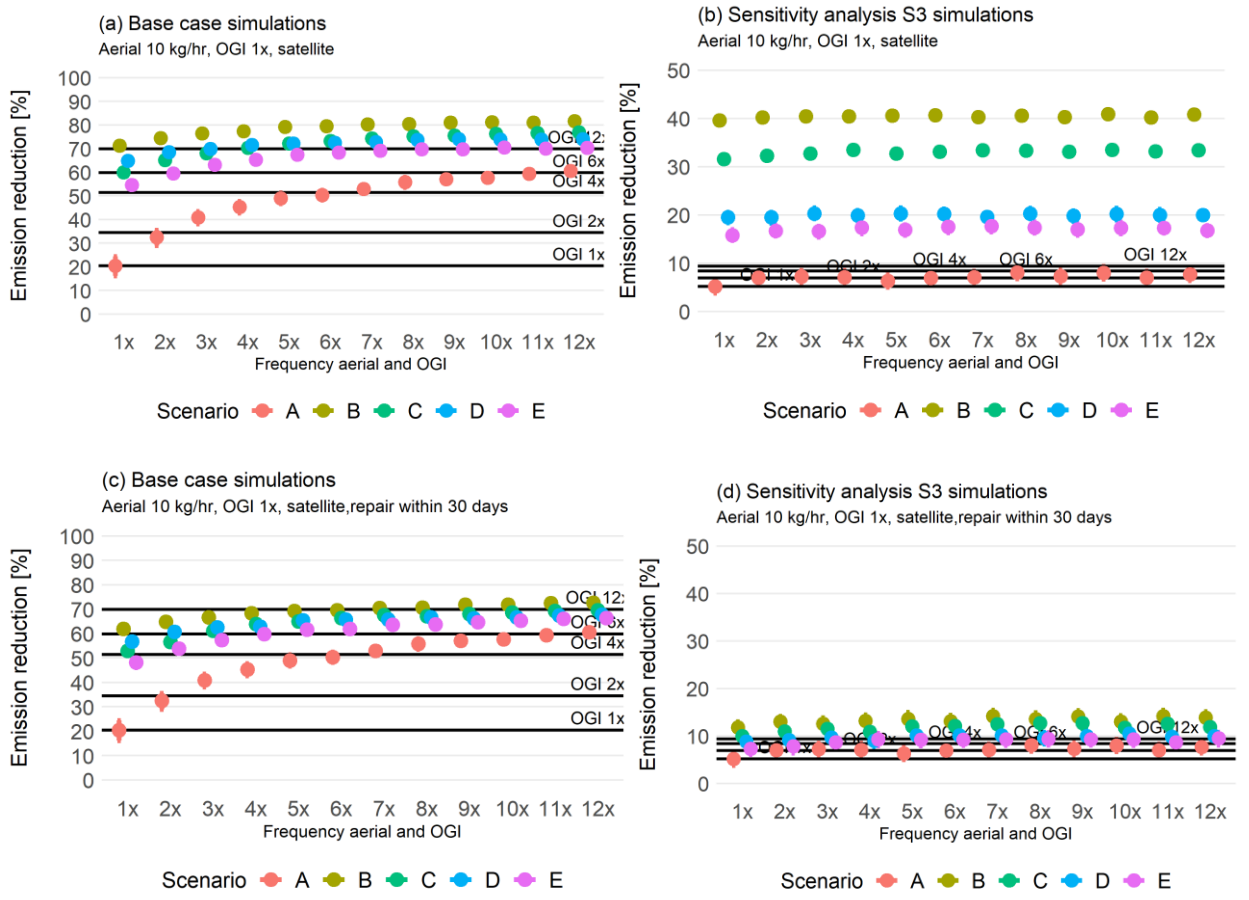


Figure S10. Emission reduction of LDAR scenarios with satellite + aerial + OGI, comparing repair times for sources detected with satellites prioritized vs. all within 30 days for (a) base case (long duration of high-emitters) and fixing emissions detected by satellites quickly, (b) sensitivity analysis S3 (short duration of high-emitters) simulations and fixing emissions detected by satellites quickly, (c) base case simulations (long duration of high-emitters) and fixing all emissions within 30 days, and (d) sensitivity analysis S3 (short duration of high-emitters) and fixing all emissions within 30 days. Scenario A is no satellite. Scenario B is satellite with 50 kg/hr and daily revisits. Scenario C is satellite with 100 kg/hr and daily revisits. Scenario D is satellite with 50 kg/hr and weekly revisits. Scenario E is satellite with 100 kg/hr and weekly revisits. Horizontal black lines indicate the reduction of OGI only LDAR programs. The horizontal axis represents the frequency of aerial and OGI, and OGI is used only once a year in all scenarios so the variation on frequency is due to the aerial surveys.

The effect of varying repair times was also assessed in scenarios with continuous monitors (Figure S11), where it was assumed that all leaks would be repaired randomly within 30 days of being found, independently of their magnitude, versus repairing those found with satellites within 1-2 days, and those with continuous monitors and above 10 kg/hr within 7 days ("Tiered repair" in Figure S11). All scenarios achieve larger reductions than 12x OGI, under both long duration of high-emitters (Figure S11a) and short duration of high-emitters (Figure S11b). However, repairing larger emissions quicker can lead to significantly more reductions; the effect is even more pronounced in scenarios with short duration of high-emitters.

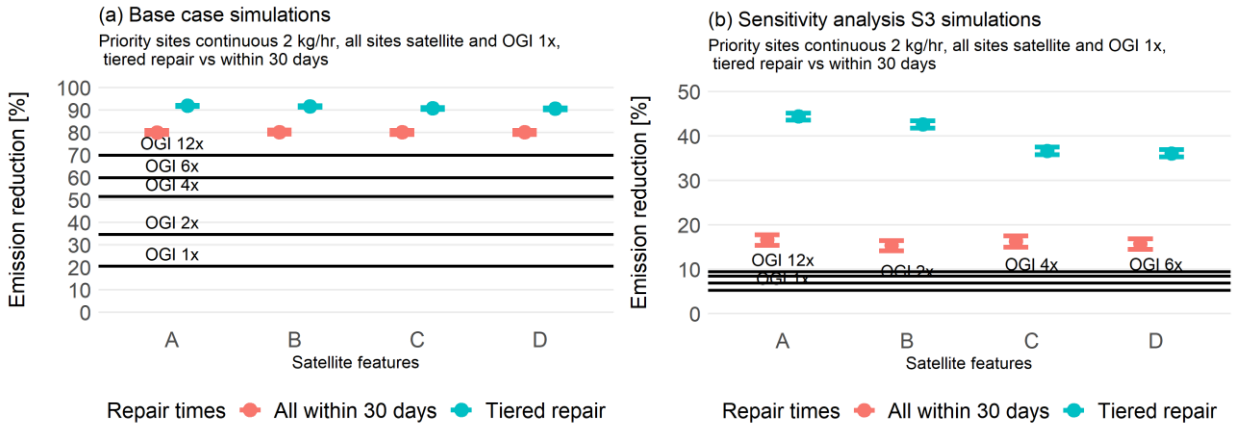


Figure S11. Emission reduction of LDAR scenarios with continuous monitors at tank batteries, satellites and OGI, comparing repair times for sources detected with satellites and continuous monitors faster vs. all repairs within 30 days for (a) base case (long duration of high-emitters) and (b) sensitivity analysis S3 (short duration of high-emitters) simulations. LDAR type A is satellite with detection threshold of 50 kg/hr and daily revisit. LDAR type B is satellite with detection threshold of 100 kg/hr and daily revisit. LDAR type C is satellite with detection threshold of 50 kg/hr and weekly revisit. LDAR type D is satellite with detection threshold of 100 kg/hr and weekly revisit. Horizontal black lines indicate the reduction of OGI only LDAR programs.

S.10. Number of LDAR Hours Required

The methodology to estimate the LDAR hours required in each program is based on Sridharan et al.[12] It was assumed that one hour of ground work is needed to inspect tank batteries and 0.1667 hours to scan wellhead sites whenever there were full OGI inspections. On follow-ups that require a survey after an initial detection from an aircraft, satellite or continuous monitor, it was assumed that the groundwork was 0.1667 for tank batteries because only a certain section of the facility was

inspected rather than the whole facility. Once ground hours were estimated, LDAR administrative hours were calculated by assuming that 80% of the LDAR work hours correspond to ground work and 20% by administrative work.[12] For full facility inspections, one hour of driving was assumed for every 2.3 hours of work (ground plus administrative hours), as described by Sridharan et al.[18] However, for follow-ups, given that the time to inspect tank batteries was decreased, 1 hour of driving was assumed for every 0.38 hours of work, as driving times to a facility are not affected whether it is a full inspection or a follow-up. The hours reported include ground, administrative and driving time. In the model only non-routine sources were simulated, thus this number of hours here reported is a lower bound, as more driving time could be spent since detection technologies capture routine sources as well.

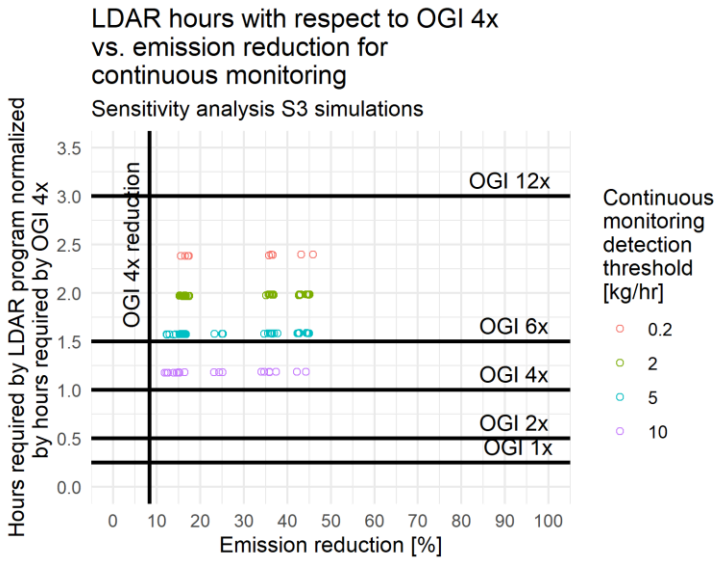


Figure S12. Number of hours required by LDAR programs normalized by the hours required by OGI 4x vs. their emission reduction, for sensitivity analysis S3 scenarios (short duration of high-emitters) that include continuous monitors. Horizontal black lines indicate the number of visits of OGI only LDAR programs. The vertical black line indicates the reduction of OGI 4x.

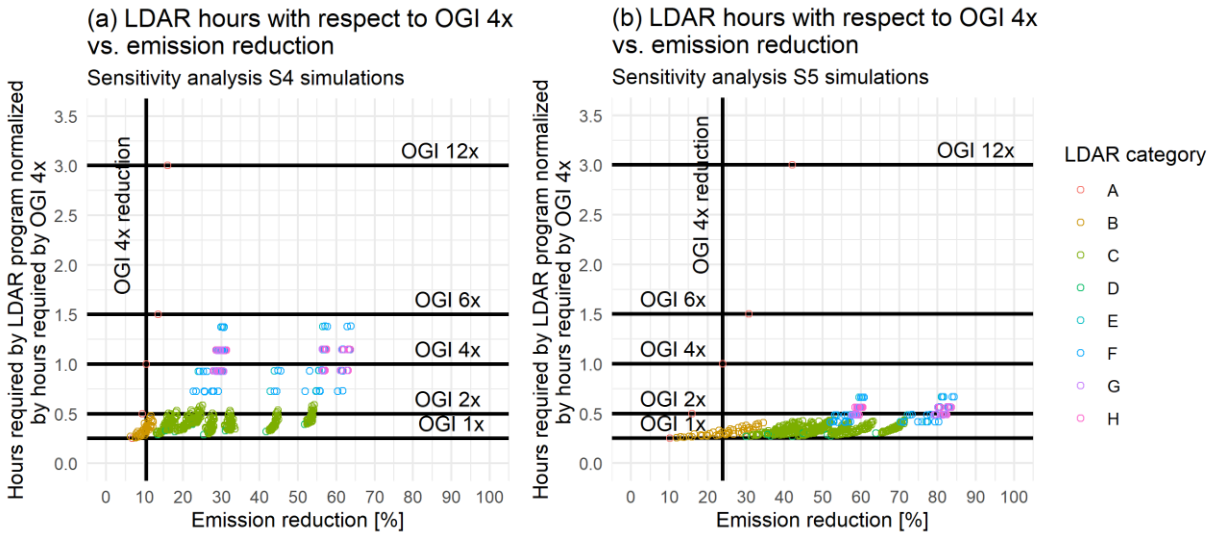
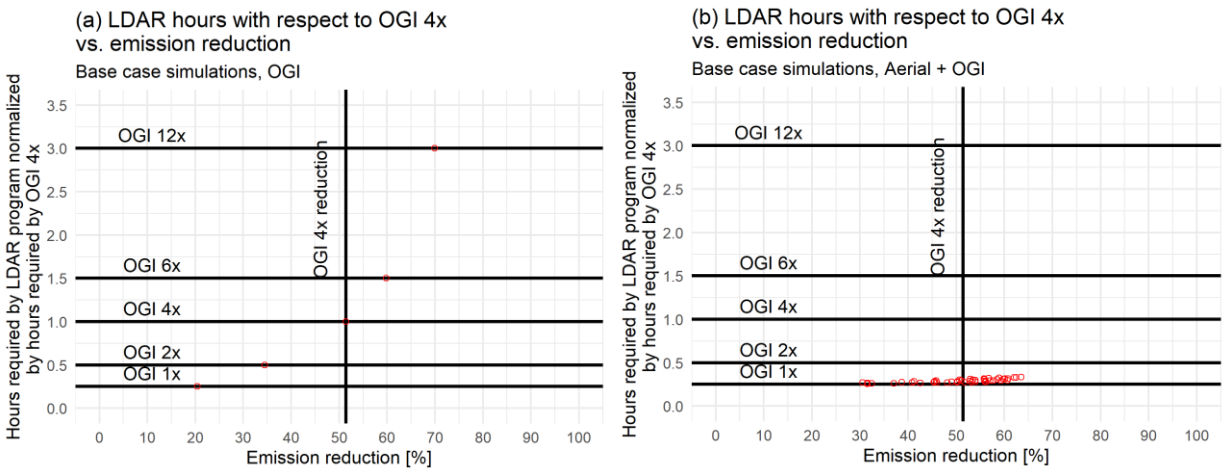


Figure S13. Number of hours required by LDAR programs normalized by the hours required by OGI 4x vs. their emission reduction for (a) sensitivity analysis S4 and (b) sensitivity analysis S5 simulations. The data points are colored by LDAR category: "A" = OGI, "B" = Aerial + OGI, "C"= Satellite + aerial + OGI, "D" = Satellite + OGI, "E" = Continuous + OGI, "F" = Satellite + continuous + OGI, "G" = Satellite + continuous at priority sites + OGI, "H" = Satellite + continuous at priority sites + aerial at non priority sites + OGI. Horizontal black lines indicate the number of visits of OGI only LDAR programs. The vertical black line indicates the reduction of OGI 4x. Each data point corresponds to one scenario in a particular LDAR category, with differences in detection threshold, frequency of inspections and repair times.



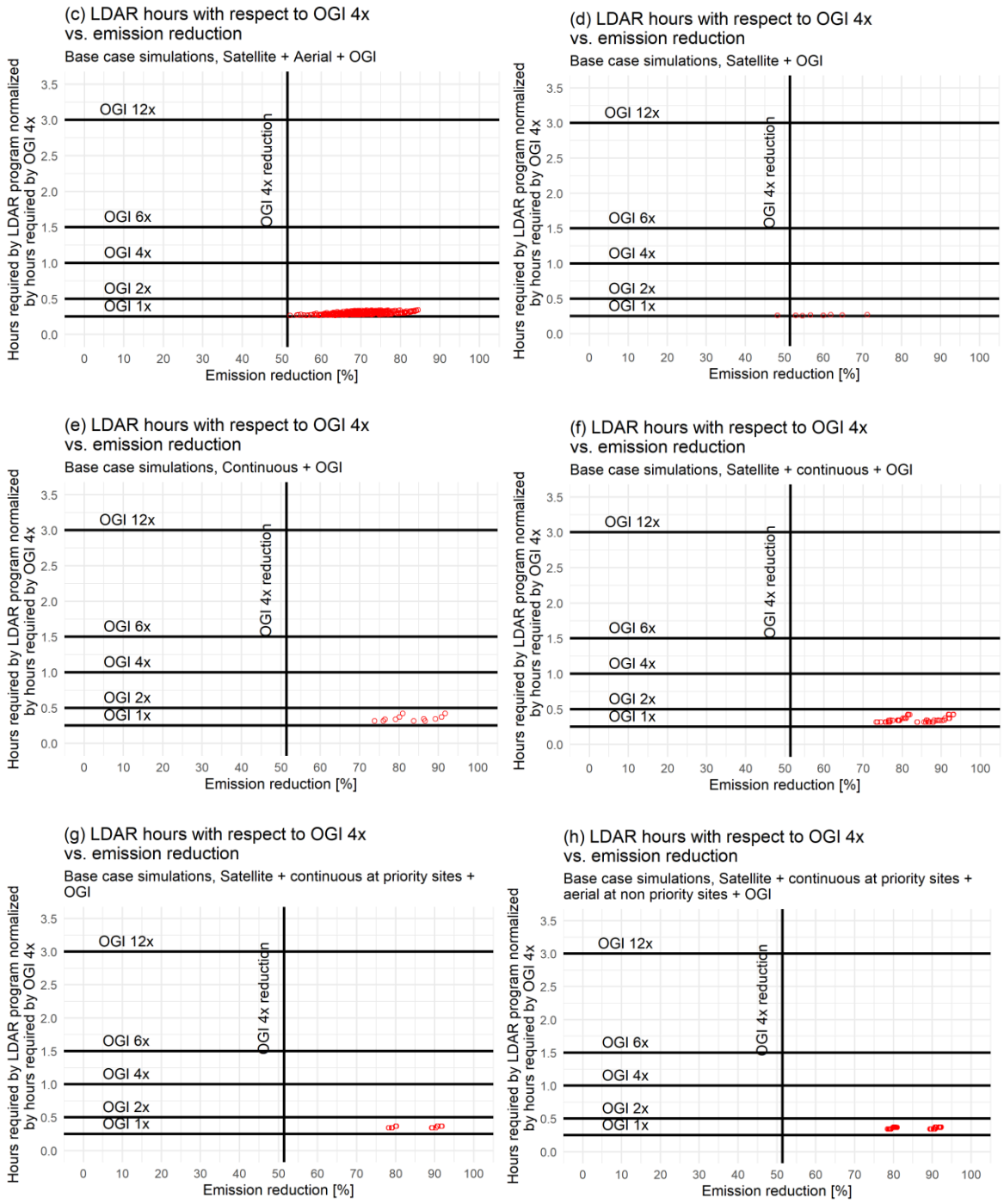
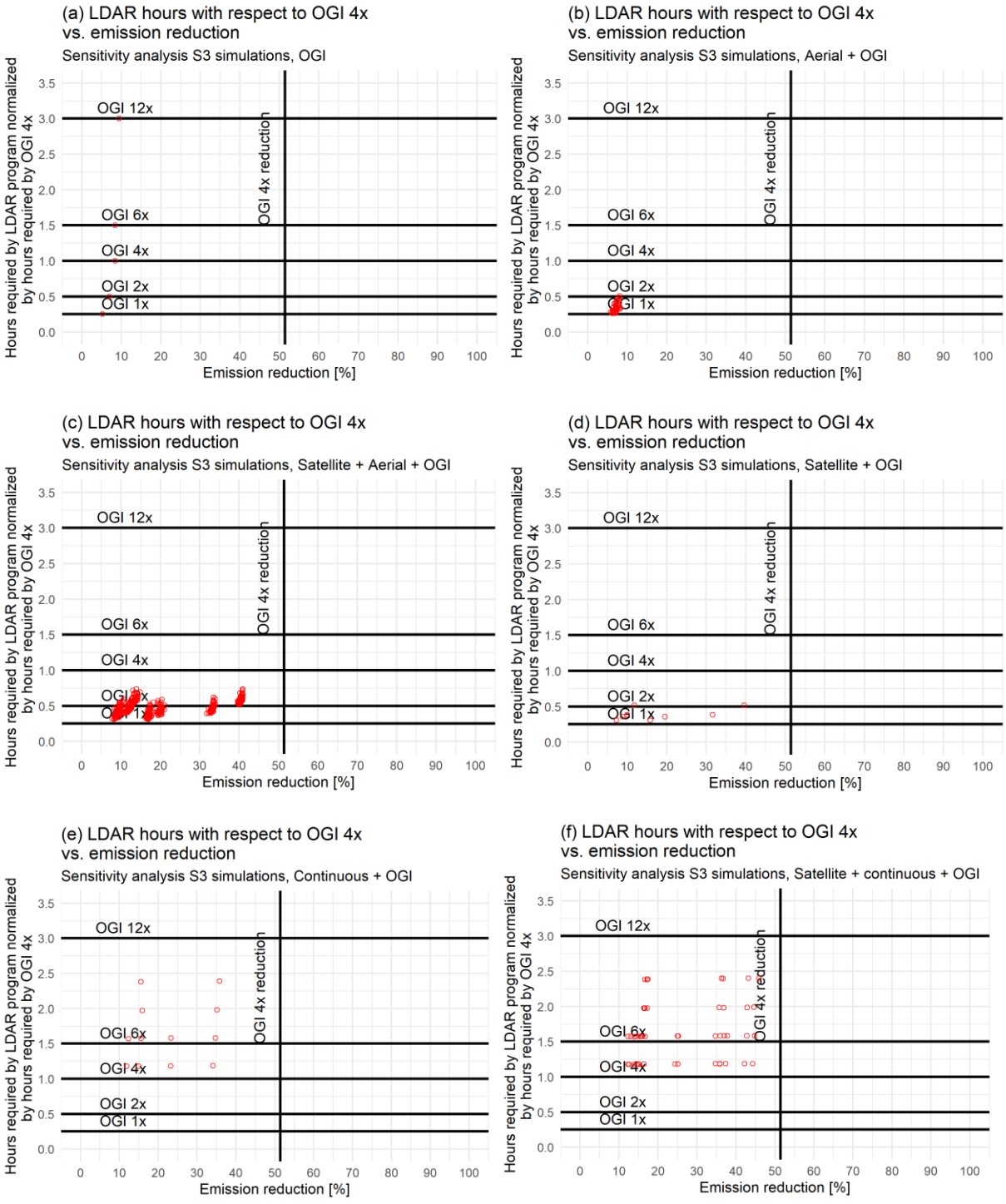


Figure S14. Number of hours required by LDAR programs normalized by the hours required by OGI 4x vs. their emission reduction for base case (long duration of high-emitters) for (a) OGI, (b) Aerial + OGI, (c) satellite + aerial + OGI, (d) satellite + OGI, (e) continuous + OGI, (f) satellite + continuous + OGI, (g) satellite + continuous at priority sites + OGI, (h) satellite + continuous at priority sites + aerial at non priority sites + OGI. Horizontal black lines indicate the number of visits of OGI only LDAR programs. The vertical black line indicates the reduction of OGI 4x. Each data point corresponds to one scenario in a particular LDAR category, with differences in detection threshold, frequency of inspections and repair times.


(a) LDAR hours with respect to OGI 4x vs. emission reduction
Sensitivity analysis S3 simulations, OGI
(b) LDAR hours with respect to OGI 4x vs. emission reduction
Sensitivity analysis S3 simulations, Aerial + OGI
(c) LDAR hours with respect to OGI 4x vs. emission reduction
Sensitivity analysis S3 simulations, Satellite + Aerial + OGI
(d) LDAR hours with respect to OGI 4x vs. emission reduction
Sensitivity analysis S3 simulations, Satellite + OGI
(e) LDAR hours with respect to OGI 4x vs. emission reduction
Sensitivity analysis S3 simulations, Continuous + OGI
(f) LDAR hours with respect to OGI 4x vs. emission reduction
Sensitivity analysis S3 simulations, Satellite + continuous + OGI
Hours required by LDAR program normalized by hours required by OGI 4x
Emission reduction [%]
OGI 12x
OGI 6x
OGI 4x
OGI 2x
OGI 1x
OGI 4x reduction

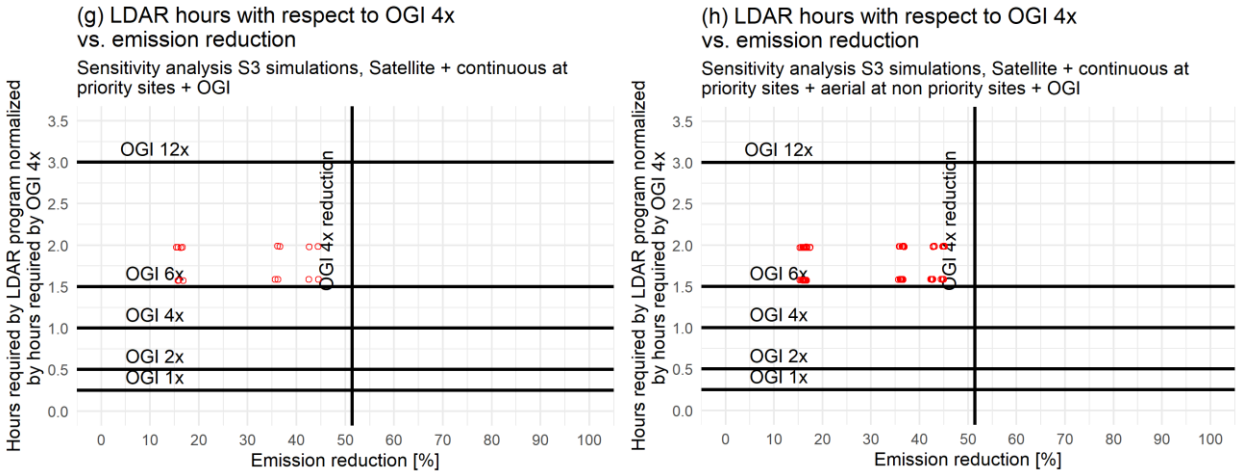


Figure S15. Number of hours required by LDAR programs normalized by the hours required by OGI 4x vs. their emission reduction for sensitivity analysis S3 simulations (short duration of high-emitters) for (a) OGI, (b) Aerial + OGI, (c) satellite + aerial + OGI, (d) satellite + OGI, (e) continuous + OGI, (f) satellite + continuous + OGI, (g) satellite + continuous at priority sites + OGI, (h) satellite + continuous at priority sites + aerial at non priority sites + OGI. Horizontal black lines indicate the number of visits of OGI only LDAR programs. The vertical black line indicates the reduction of OGI 4x. Each data point corresponds to one scenario in a particular LDAR category, with differences in detection threshold, frequency of inspections and repair times.

S.11. Effect of Time to Leak Recurrence

In order to assess the effect that leak recurrence might have on emission reduction, simulations were ran assuming that once a leak has been repaired for a particular source, it will take 10 times longer for that particular source to leak again. Thus, from Eq. 3, MTTR is kept constant, while MTBF is increased by a factor of 10 once a source has been repaired as part of an LDAR program. The simulations were performed only for the scenario of satellite with daily revisit and 50 kg/hr detection threshold, plus aerial surveys with 10 kg/hr detection threshold and yearly OGI. As shown in Figure S16, in scenarios where it takes more time for particular sources to leak again, the emission reduction across all LDAR scenarios increases, with a more pronounced effect in the scenarios with short duration of high-emitters (sensitivity analysis S3). More data of leak recurrence is needed to have more accurate predictions of absolute reduction achieved by LDAR programs.

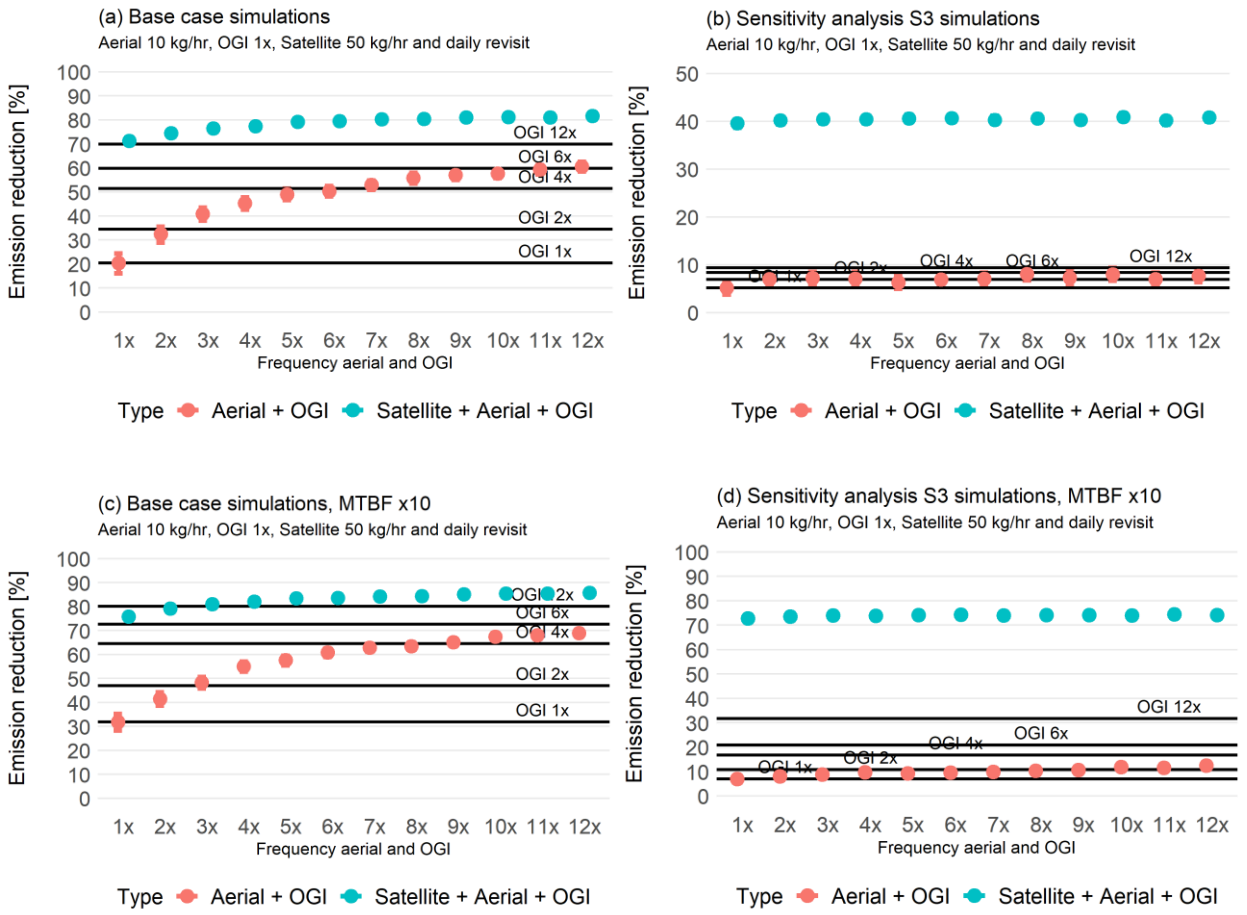


Figure S16. Emission reduction of LDAR scenarios with aerial + OGI and satellite + aerial + OGI for (a) base case simulations (long duration of high-emitters) and assuming same time for recurrence of a leak, (b) sensitivity analysis S3 (short duration of high-emitters) and assuming same time for recurrence of a leak, (c) base case simulations (long duration of high-emitters) and assuming 10 times longer for leak recurrence once a source has been fixed in an LDAR program, and (d) sensitivity analysis S3 (short duration of high-emitters) and assuming 10 times longer for leak recurrence once a source has been fixed in an LDAR program. The horizontal axis represents the frequency of aerial and OGI, and OGI is used only once a year in all scenarios so the variation on frequency is due to the aerial surveys. Horizontal black lines indicate the number of visits of OGI only LDAR programs.

S.12. References

1. Stokes, S., Tullos, E., Morris, L., Cardoso-Saldaña, F. J., Smith, M., Conley, S., ... & Allen, D. T. (2022). Reconciling Multiple Methane Detection and Quantification Systems at Oil and Gas Tank Battery Sites. *Environmental Science & Technology*, 56(22), 16055-16061.
2. Ravikumar, A.P. (2022). FEAST: US – Alternative LDAR programs for representative US O&G Production Facilities. Comment to US Environmental Protection Agency submitted by Environmental Defense Fund, Attachment L – FEAST National Slides. https://www.regulations.gov/comment/EPA-HQ-OAR-2021-0317-0844 (Accessed August 1, 2022).
3. Rutherford, J. S., Sherwin, E. D., Ravikumar, A. P., Heath, G. A., Englander, J., Cooley, D., ... & Brandt, A. R. (2021). Closing the methane gap in US oil and natural gas production emissions inventories. *Nature communications*, *12*(1), 1-12.
4. U.S. Environmental Protection Agency. Greenhouse Gas Reporting Program. https://www.epa.gov/ghgreporting (Accessed August 1, 2022).
5. Chen, Y., Sherwin, E. D., Berman, E. S., Jones, B. B., Gordon, M. P., Wetherley, E. B., ... & Brandt, A. R. (2022). Quantifying Regional Methane Emissions in the New Mexico Permian Basin with a Comprehensive Aerial Survey. *Environmental Science & Technology*, *56*(7), 4317-4323.
6. Cusworth, D. H., Duren, R. M., Thorpe, A. K., Olson-Duvall, W., Heckler, J., Chapman, J. W., ... & Miller, C. E. (2021). Intermittency of large methane emitters in the Permian Basin. *Environmental Science & Technology Letters*, *8*(7), 567-573.
7. Tyner, D. R., & Johnson, M. R. (2021). Where the Methane Is—Insights from Novel Airborne LiDAR Measurements Combined with Ground Survey Data. *Environmental Science & Technology*, *55*(14), 9773-9783.
8. Kemp, C. E., Ravikumar, A. P., & Brandt, A. R. (2016). Comparing natural gas leakage detection technologies using an open-source "virtual gas field" simulator. *Environmental science & technology*, *50*(8), 4546-4553.
9. Fox, T. A., Gao, M., Barchyn, T. E., Jamin, Y. L., & Hugenholtz, C. H. (2021). An agent-based model for estimating emissions reduction equivalence among leak detection and repair programs. *Journal of Cleaner Production*, *282*, 125237.
10. National Centers for Environmental Information, US National Ocean and Atmospheric Administration (NOAA). (2018). Comparative Climatic Data (CCD). https://www.ncei.noaa.gov/products/land-based-station/comparative-climatic-data (Accessed August 1, 2022).
11. Zimmerle, D., Vaughn, T., Bell, C., Bennett, K., Deshmukh, P., & Thoma, E. (2020). Detection limits of optical gas imaging for natural gas leak detection in realistic controlled conditions. *Environmental science & technology*, 54(18), 11506-11514.
12. Sridharan, S., Lazarus, A., Reese, C., Wetherley, E., Bushko, K., & Berman, E. (2020). Long Term, Periodic Aerial Surveys Cost Effectively Mitigate Methane Emissions. In *SPE Annual Technical Conference and Exhibition*. OnePetro.