Beyond Snapshots: Building Methane Measurement-Informed Inventories for Midstream Oil and Gas Sites in the Appalachian Basin

Arthur Santos[a,], Jacob Mdigo[a], Anna Hodshire[a], Daniel Zimmerle[a], Tecle Rufael[b], Matthew R. Harrison[b], Arvind Ravikumar[c]

[a] *Energy Institute, Colorado State University, Fort Collins, 80523, CO, USA*
[b] *SLR Consulting, Houston, 77042, TX, USA*
[c] *Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, 78712, TX USA*

S1. Extracting Failure Statistics from Measurements

The Mechanistic Air Emissions Simulator (MAES) simulations were designed to include emissions from upset conditions using two approaches:

- **Traditional method**: Emissions were estimated using the product of activity factor (AF) and emission factor (EF). This approach was applied to compressor rod-packing large emitters, tank emissions, and facility piping leaks. Failure frequency, duration, and emission size for each type were calculated based on aerial campaign observations.
- **Mechanistic modeling**: Emissions were estimated based on mechanistic processes at the sites, accounting for equipment states and fluid flows. This approach was applied to heaters and flares, where measurements defined the fraction of time equipment spent in each state, and emission size and duration were estimated based on those states and fluid flows.

Details on failure types and the two modeling approaches are provided by Santos et al. [1]. The following subsections detail how aerial measurement data informed estimates of emissions frequency and duration for upset conditions in flares, heaters, compressors, tanks, and facility piping.

S1.1. Failure Frequencies

Detections were first screened against the facility cumulative distribution function (CDF) from the Inventory model, and detections coinciding with logged maintenance were excluded. For the remaining detections, operator cause analyses were requested and were used both to exclude events and to inform source attribution. Each detection retained as an abnormal process event carries a classification confidence weight $P_{i,j}$ reflecting the strength of the evidence supporting its classification. Confidence weights were drawn from a fixed set of levels rather than assigned on a continuous scale. Table S2 gives the levels, the basis for each level, and the number of events at each. Where unit specific evidence was available, from an operator cause

Email address: arthur.santos@colostate.edu ()

Table S1: Equipment count per site, total equipment count across all sites, and total equipment scanned. Total equipment scanned was calculated by multiplying the total equipment count by 3 to account for measurement campaigns in Q2, Q3, and Q4 of 2023.

Facility	Separators	Compressors	Oil Tank Battery	Water Tank Battery	Heaters	Flares	Dehydrators	Facility Piping
A	13	7	2	2	7	2	3	1
B	12	4	0	2	3	2	2	1
C	12	8	8	0	4	8	1	1
D	2	1	0	1	1	0	0	1
E	3	3	0	1	1	0	0	1
F	6	2	0	2	1	2	2	1
G	6	5	0	2	5	1	1	1
H	5	2	2	0	7	2	1	1
I	3	1	0	1	1	1	1	1
J	6	1	0	2	3	1	1	1
K	3	1	0	2	0	1	1	1
L	2	0	0	2	1	3	3	1
M	5	0	0	1	0	1	1	1
N	3	0	0	2	1	2	2	1
O	1	3	0	1	1	2	2	1
P	4	2	1	0	7	1	1	1
Q	6	0	2	1	12	2	2	1
R	11	0	0	2	2	3	3	1
S	2	0	0	1	1	1	1	1
Total Equip.	**105**	**40**	**15**	**25**	**58**	**35**	**28**	**19**
Total Equip. Scanned	**315**	**120**	**45**	**75**	**174**	**105**	**84**	**57**

analysis, from plume location and morphology, or from a measured rate that could not be reconciled with the modeled normal operation range for that unit, the weight reflects that evidence directly. Where no discriminating evidence was available, the weight reflects the general plausibility of an abnormal cause for that detection. In a small number of cases an abnormal classification was retained despite an operator report of normal operation, where the measured magnitude could not be explained by any normal operation source at that unit. All classifications were performed by the first author and were not independently re-coded. The probability weighted failure counts in Table 2 follow from these weights through Equation 7.

Table S2: Classification confidence scale applied to detections retained as abnormal process events. The confidence weight $P_{i,j}$ enters the probability weighted failure count through Equation 7. Counts are pooled across all equipment categories, operators, and survey rounds.

P	**Basis for assignment**	**Events**
1.00	Operator cause analysis identified an abnormal condition, or the measured rate exceeded the modeled normal operation range for that unit by a margin that no normal operation source could account for.	13
0.75	Source attribution was clear and an abnormal or unreported source was the most likely explanation, but no unit specific evidence was available to confirm it.	48
0.50	An abnormal cause and a normal operation cause were judged about equally plausible, or a credible measurement or attribution artifact could not be excluded.	13
0.20–0.30	Plume imagery or site context did not clearly support the assigned source. Events were retained at low weight rather than discarded.	5
Total		**79**

Because the study involves only two operators, disclosing event level detection, investigation, and exclusion counts by equipment category, together with the site level equipment counts in Table S1, could allow re-identification of individual operator assets and reverse engineering of the confidential, operator specific investigation thresholds. We therefore report the classification confidence scale in Table S2 together with the units scanned, the probability weighted failure count, and the resulting failure rate per equipment category, which are the quantities that enter the model and are given in Table 2 of the main manuscript, and we withhold the intermediate event level breakdown under the coalition data sharing agreement.

S1.2. Simulated Failure Types and their Durations

S1.2.1. Flares

Flares, Enclosed Combustion Device (ECD)s, and Vapor Destruction Unit (VDU)s are all simulated the same way, and they can operate in three states: operating, malfunctioning, and unlit, with each state affecting operating efficiency and emissions. In the unlit state, the flare releases all gas input directly to the atmosphere. The destruction efficiencies for the operating, malfunctioning, and unlit states are assumed to be 98%, 90%, and 0%, respectively.

Once in the operating state, the equipment can transition to malfunctioning or unlit states based on their respective probabilities (pMalfunction and pUnlit). In the malfunctioning or unlit state, the equipment can only transition back to the operating state with a probability of 1. MAES requires these probabilities as input to estimate the steady-state vector for each state. Transition probabilities are calculated using observed state fraction of time from aerial surveys, supplemented by cause analyses from operators, and applying a Markov transition methodology [2].

During the pilot campaign, flares were observed in the malfunctioning state for 3.8% of the time. No unlit flares were detected, but a fraction of time of 0.4% was assumed for this state due to the limited number of surveys. The remaining fraction of time in the operating state was 95.8%. Since measurement surveys were conducted quarterly, we assumed a malfunctioning state could last between 1 and 90 days, whereas an unlit state, being easier to detect, was assumed to last between 1 and 5 days.

Using the script in Section 26.3.2 of the MAES Study Sheet Guide [2], the Flare Transition Matrix (Q) was calculated from the Fraction of Time in States (F) matrix. The resulting pMalfunction and pUnlit probabilities for the pilot study were 0.385 and 0.615, respectively.

S1.2.2. Heaters

Similar to flares, heaters are simulated in three states: operating, malfunctioning, and shut-in. Heaters are essential in separators to heat fluids, aiding the separation process and preventing wax formation. However, during warm weather, heaters may not be needed, entering a shut-in state with zero emissions. In the operating and malfunctioning states, the heater operates with an assumed destruction efficiencies of 99% and 80%, respectively.

Once in the operating state, heaters can transition to malfunctioning or shut-in states based on their respective probabilities (pMalfunction and pShutIn). In the malfunctioning or shut-in states, heaters can only transition back to the operating state. Transition probabilities are calculated based on aerial survey data, supplemented by operator-provided cause analyses, and determined using the Markov transition methodology as well.

During the pilot campaign, emissions coming from separators were classified as malfunctioning heater, as all detected emissions for this equipment category were above 3.7 kg/h. As the Appalachian basin has a temperate climate with cold winters, the heaters were assumed to be in the Shut In state 50% of the time,

since they operate day and night during the winter (100% of the time in 3 months) and half of Spring/Fall during colder nights (50% of the time in 6 months). From the measurement campaign, heaters were observed in the Malfunctioning state 9.5% of the time, consistent with the failure rate reported in Table 2. The remaining fraction of time in the operating state was 40.5%. Since measurement surveys were conducted quarterly, we assumed a malfunctioning state could last between 1 and 90 days, whereas the shut in state was assigned a duration of 183 days.

The script in Section 26.3.2 of the MAES Study Sheet Guide [2] was used to calculate the Heater Transition Matrix (Q) from a Fraction of Time in States (F) matrix. The calculated pMalfunction and pShutIn for the pilot study were 0.432 and 0.568, respectively.

S1.2.3. Compressors

The failure event simulated for compressors is seal failure. In natural gas compressors, seals are critical components that prevent gas leakage where rotating or moving parts, like shafts or rods, pass through the compressor casing. In reciprocating compressors, for instance, the rod packing seal minimizes gas escape around the piston rod. In centrifugal compressors, the shaft seal prevents leakage around the rotating shaft. When these seals wear down, they lose their ability to tightly contain pressurized gas, leading to increased emissions in these equipment.

MAES estimates seal emissions in two categories: normal seal vents and large emitter seal vents. The normal category represents emissions during regular compressor operations, while large emitters represent seal emissions when seals are worn or failing. However, this model in MAES needs improvement, as currently both states can occur simultaneously, which does not reflect real-world conditions. To correct this, post-processing was applied to remove normal seal emissions whenever a large emitter event occurred.

Normal seal vent emissions are set based on EF from the GEF study [3], a national campaign that measured emissions at gathering and boosting facilities across the U.S. Large emitter seal vent emissions were determined using an emission factor distribution derived from aerial measurements, supplemented by operator-provided cause analysis and evaluation by the study team. This data also enabled the estimation of event frequency, calculated as the number of detected large emitter events relative to the total number of compressors measured. The emission size for these large emitter events was drawn from the measurement derived emission factor distribution, following the post processing described in Section 2.3.3.

S1.2.4. Tanks

A similar approach to compressors was used for simulating upset conditions at tanks. While tank emissions primarily arise from open thief hatches and overpressure events, often caused by upstream stuck dump valves, a single emission factor distribution was created for tanks, following the methodology described in 2.3.3. These emissions were grouped and simulated under the "Thief Hatch" category on MAES. All other tank emissions were disabled in the simulation, except for component leaks. The event frequency for tanks was computed as the number of detected tank events divided by the total number of tank batteries scanned, and the emission

size was drawn from the emission factor distribution built from the post processed Bridger measurements described in Section 2.3.3.

S1.2.5. Facility Piping

Emissions categorized as "Facility Piping" or "Other" by Bridger were grouped and used to develop an EF distribution for Facility Piping, using the methodology described in 2.3.3. These emissions were assigned a probability and incorporated into the simulation.

S2. Scope of the Reported Uncertainty

The intervals shown on the MAES estimates in Figure 2 span the 2.5th to 97.5th percentiles of the simulated annual totals across Monte Carlo realizations. They describe the year-to-year variability the model produces, principally from stochastic sampling of equipment states and from the propagated Bridger measurement uncertainty described in Section 2.3.3. They are not confidence intervals on the mean, which are far narrower, and they do not represent the total uncertainty of the Measurement-Informed Inventory (MII) estimate. The remainder of this section separates the uncertainty terms that are propagated into those intervals from those that are not.

Four sources of variability enter these intervals. The first is the Bridger measurement uncertainty, which reaches the models through the emission factor files, where each measurement is replaced by 1,000 draws from the relative error distribution of its rate subgroup (Section 2.3.3). The second is stochastic state sampling, since equipment state occupancy and event timing are drawn afresh in every realization. The third is event duration. Each failure event takes its duration from the assigned range, 1 to 90 days for the large emitter and malfunction modes and 1 to 5 days for the unlit and vent modes. The fourth is Monte Carlo sampling error, which is 1.6% of the all sites annual total at 100 iterations (Section 2.3.4).

Several other terms fall outside these intervals, and we give the direction of each where it can be established. Two of them act to lower the estimate. Emissions below the survey detection threshold, and any event occurring between surveys, are absent from the failure statistics. Emergency shutdowns and maintenance blowdowns are not implemented in the tool, and the flare gap bounds their likely magnitude at roughly 4.8% of the MII total. The classification confidence weights work in the same direction: the 79 retained detections give a probability weighted count of 56.75, which is 28% below the unweighted count, so the weighting lowers the estimate rather than raising it. Four further terms have no direction we can establish. Both the failure counts and the scan counts are exact. The scan counts range from 57 for facility piping to 174 for heaters, each being the number of units of that type multiplied by the three survey rounds. What is not exact is how closely a rate derived from that many observations matches the long run failure probability. This statistical uncertainty is not carried through the simulation. The duration ranges are tied to the survey interval rather than assumed freely: the 90 day upper bound matches the quarterly spacing of the surveys, which is the finest resolution at which the persistence of an event could be observed. Alternative bounds were not tested.

The conversion of quarterly snapshots into annual estimates through the ergodic assumption has not been validated against continuous monitoring. Detections coinciding with logged maintenance are excluded while the surveys overlap weekday maintenance windows, and these two effects work against each other by an amount we do not quantify.

References

[1] A. Santos, W. Mollel, G. P. Duggan, A. Hodshire, P. Vora, D. Zimmerle, Using measurement-informed inventory to assess emissions in the denver-julesburg basin, ACS ES&T Air (2024).

[2] W. Mollel, J. Mdigo, A. Santos, P. Vora, J. Duggan, D. Zimmerle, Maes study sheet guide (2024).

[3] D. Zimmerle, T. Vaughn, B. Luck, T. Lauderdale, K. Keen, M. Harrison, A. Marchese, L. Williams, D. Allen, Methane emissions from gathering compressor stations in the us, Environmental Science & Technology 54 (12) (2020) 7552–7561.