COBE Statistical Model

Callan Okenberg

20 October 2025
EEMDL Annual Meeting

Introduction/Overview

High Level Model Overview

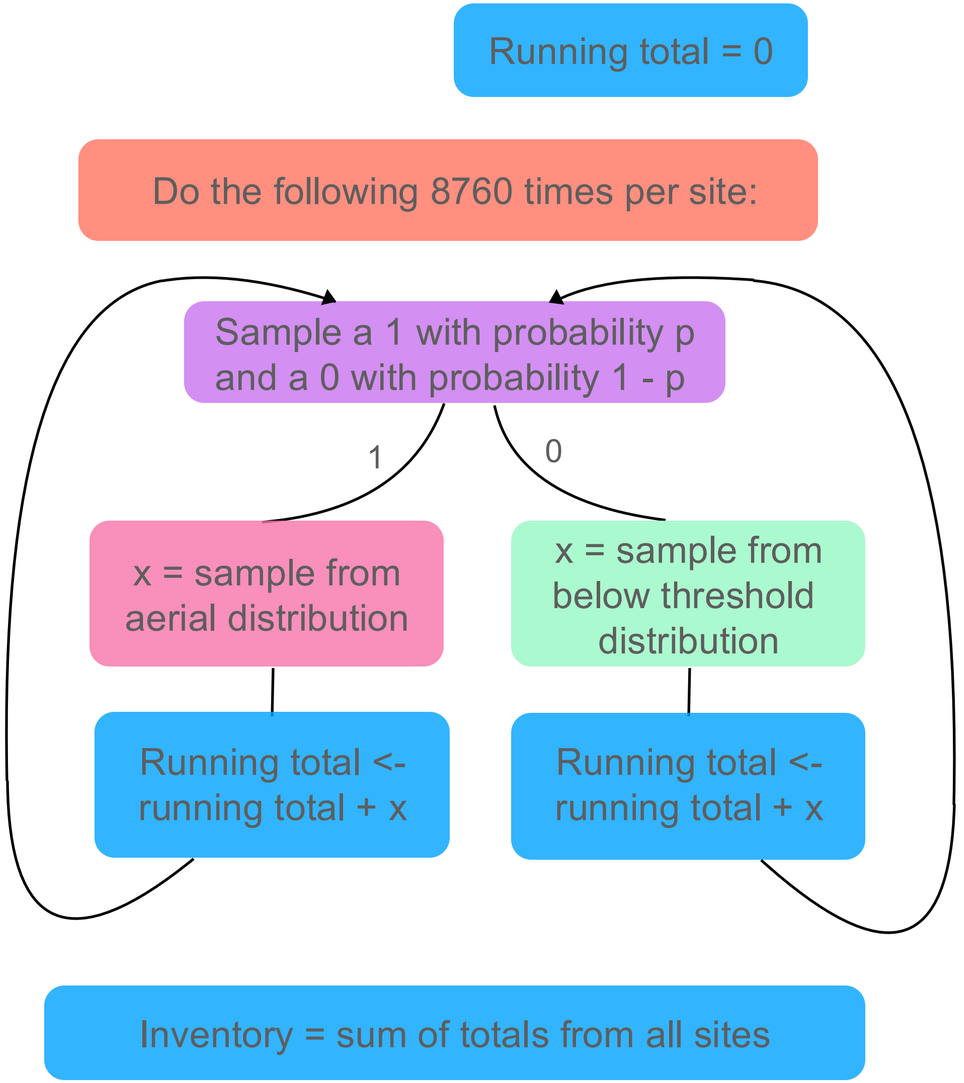
- Goal: Create a **data-driven** inventory model for the state of Colorado
- General approach is *distributional*:
  - Estimate a facility-level emissions distribution representative of production sites in Colorado
  - Repeatedly sample from this distribution to aggregate to an annual state-level inventory
- Complication: aerial technologies cannot reliably detect emissions below some threshold
  - Idea: estimate two separate distributions:
    - One from aerial data that is valid down to that threshold
    - Another that represents emissions below that threshold

Below-threshold Distributions

- Where do these below-threshold distributions come from?
- We experiment with 3 different sources:
  - Continuous Monitoring System (CMS) Data
  - 2 distributions from the literature:
    - Williams et al. (2025): assimilates many different low detection threshold technologies
    - Sherwin et al. (2024): uses the bottom-up simulation framework from Rutherford et al. (2021)

Aggregation Method

- Inputs:
  - Aerial distribution
  - Below-threshold distribution (3 different methods)
  - Probability of an emission above threshold occurring (p)
    - Estimated from the aerial data

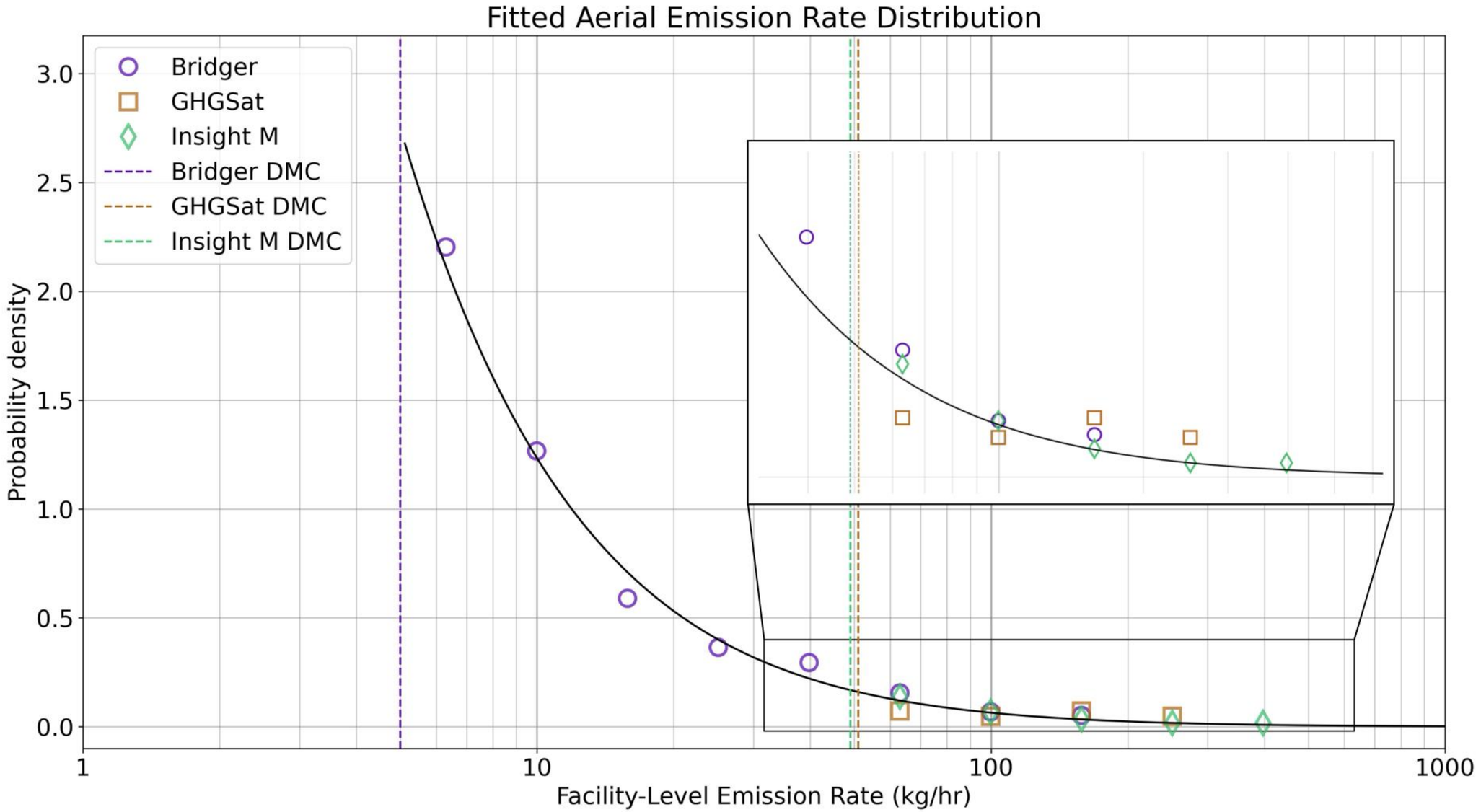


Aerial Distribution Fitting

Aerial Distribution Approach

- Goal: Combine data from three aerial vendors with different detection thresholds into one representative emissions distribution
    - Assumption: all vendors sample from the same distribution, just with different sensitivities/detection thresholds
- Conceptual approach:
    - Find a Distribution Matching Cutoff (DMC) for each vendor above which we think they can detect all emissions
        - Note: this is **not** a detection limit, rather a tool to help us combine data across vendors
    - Fit a distribution to each vendor's data **only above this DMC**
- Results in a site-level emissions distribution that is valid down to the lowest DMC

Aerial Distribution Results

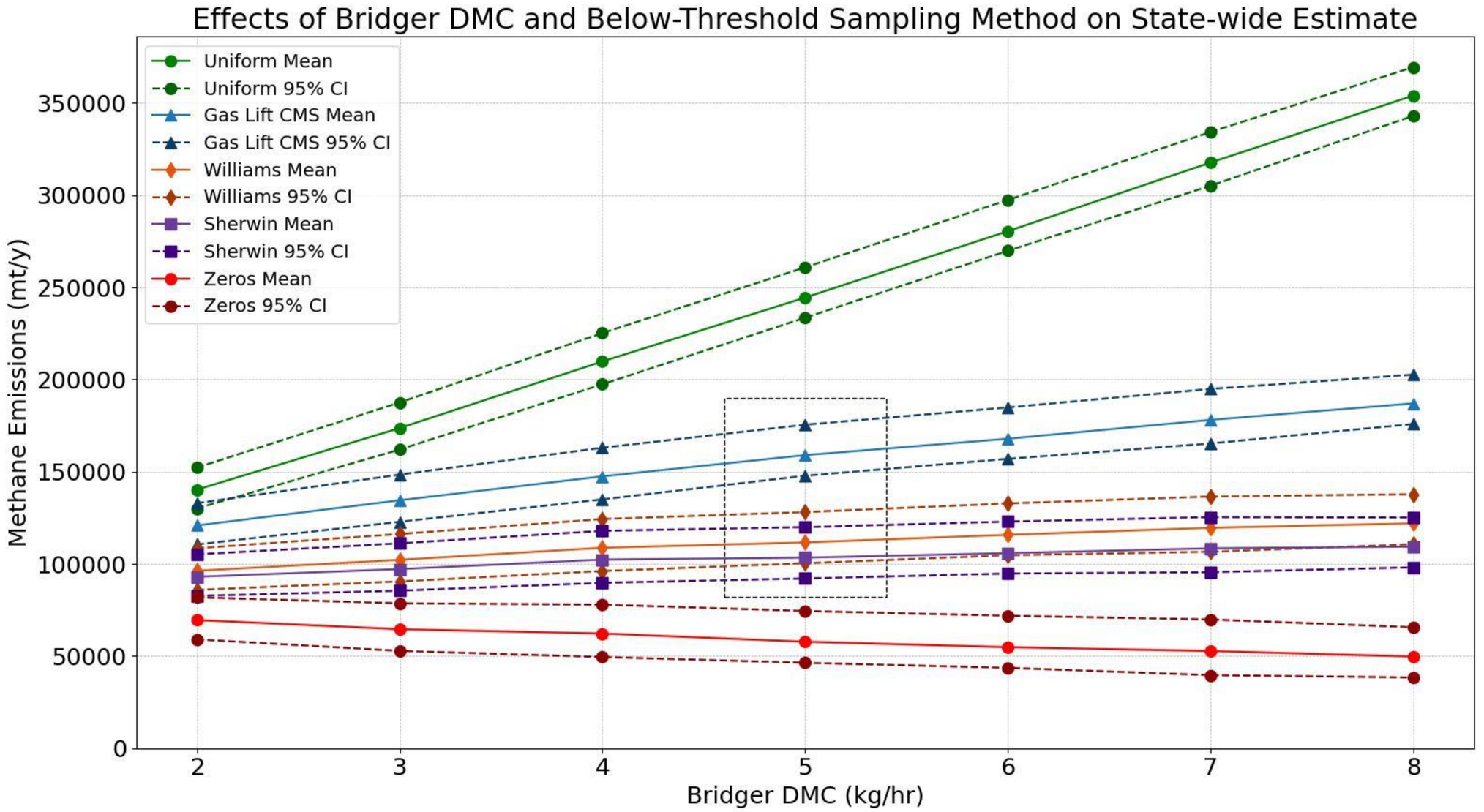


Drawbacks

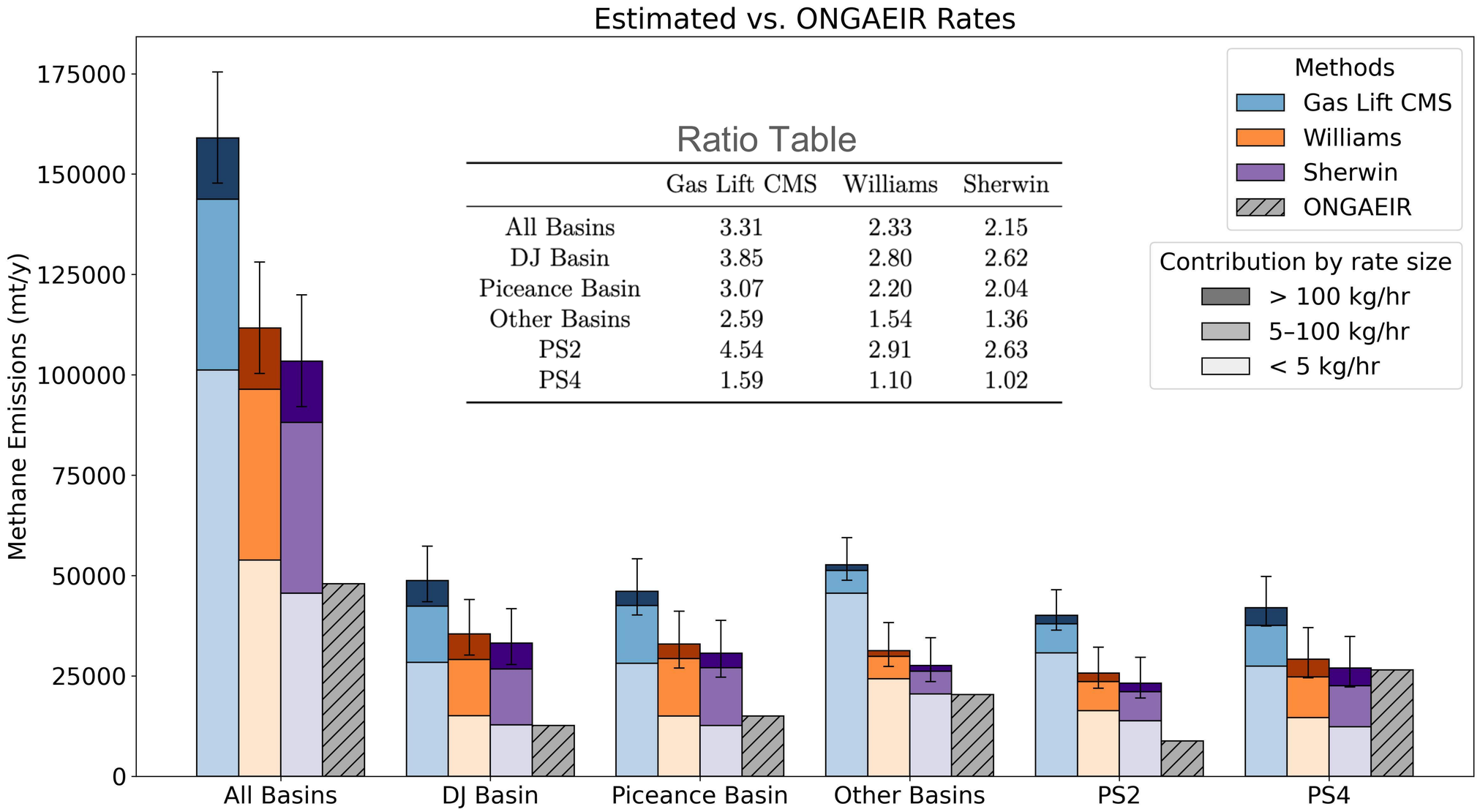
- This estimates the desired distribution, but with a few key drawbacks:
  - Two upper DMCs can be estimated by comparing how their distributions align with Bridger's above different DMCs
  - Bridger's DMC is more difficult to determine: no reference distribution to compare to
    - We estimate as 5 kg/hr, and show how sensitive our method is to different reasonable choices of Bridger DMC
  - Data below each vendor's DMC is not used

Results

Bridger DMC Sensitivity Study



State/Basin-Level Results



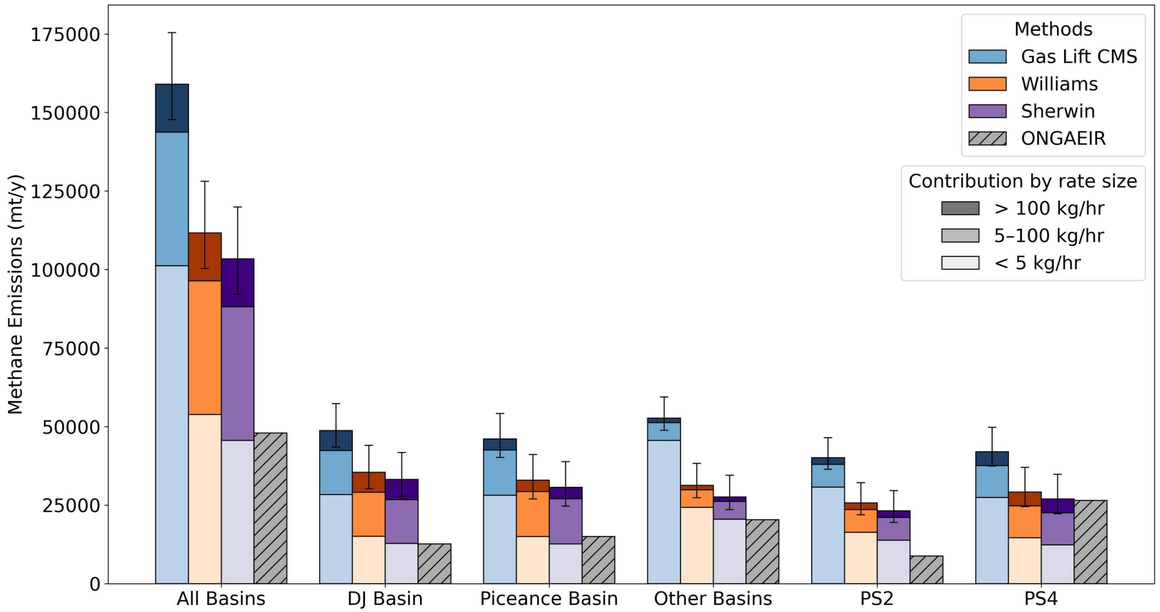
Ratio Table

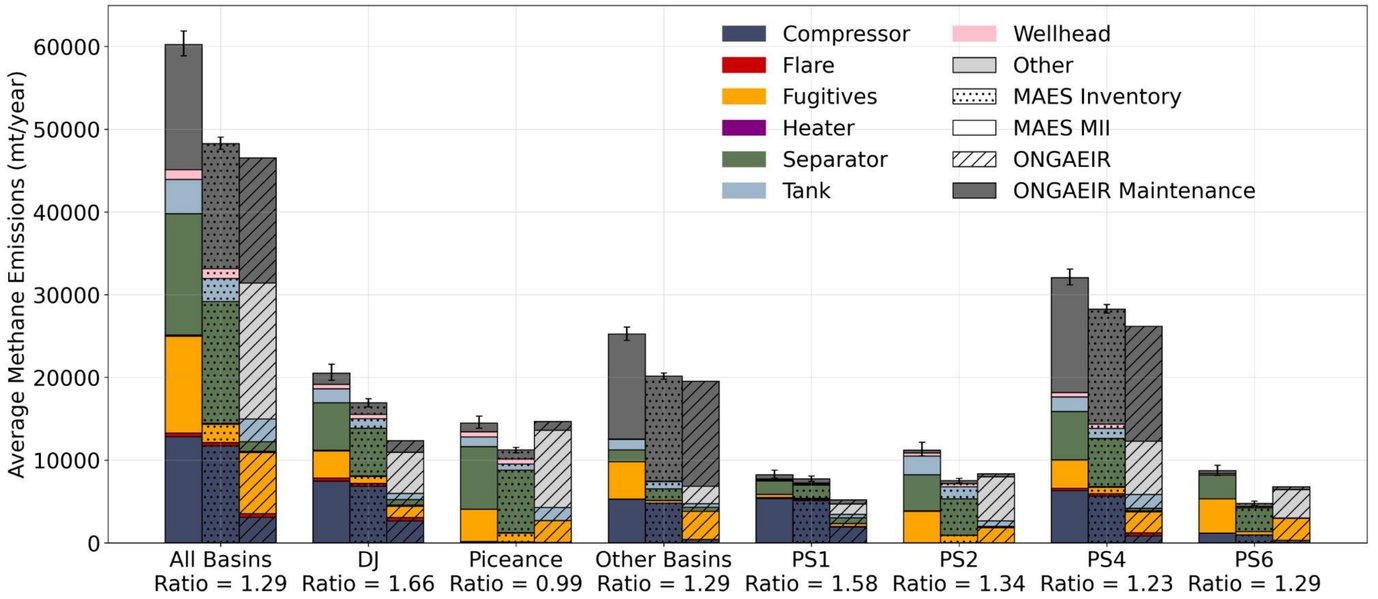
	Gas Lift CMS	Williams	Sherwin
All Basins	3.31	2.33	2.15
DJ Basin	3.85	2.80	2.62
Piceance Basin	3.07	2.20	2.04
Other Basins	2.59	1.54	1.36
PS2	4.54	2.91	2.63
PS4	1.59	1.10	1.02

MAES Comparison

Ratio Comparison

Region	Statistical Model	MAES Model
State	2.15 - 3.31	1.29
DJ Basin	2.62 - 3.85	1.66
Piceance	2.04 - 3.07	0.99
Other Basins	1.36 - 2.59	1.29
PS2	2.63 - 4.54	1.34
PS4	1.02 - 1.59	1.23



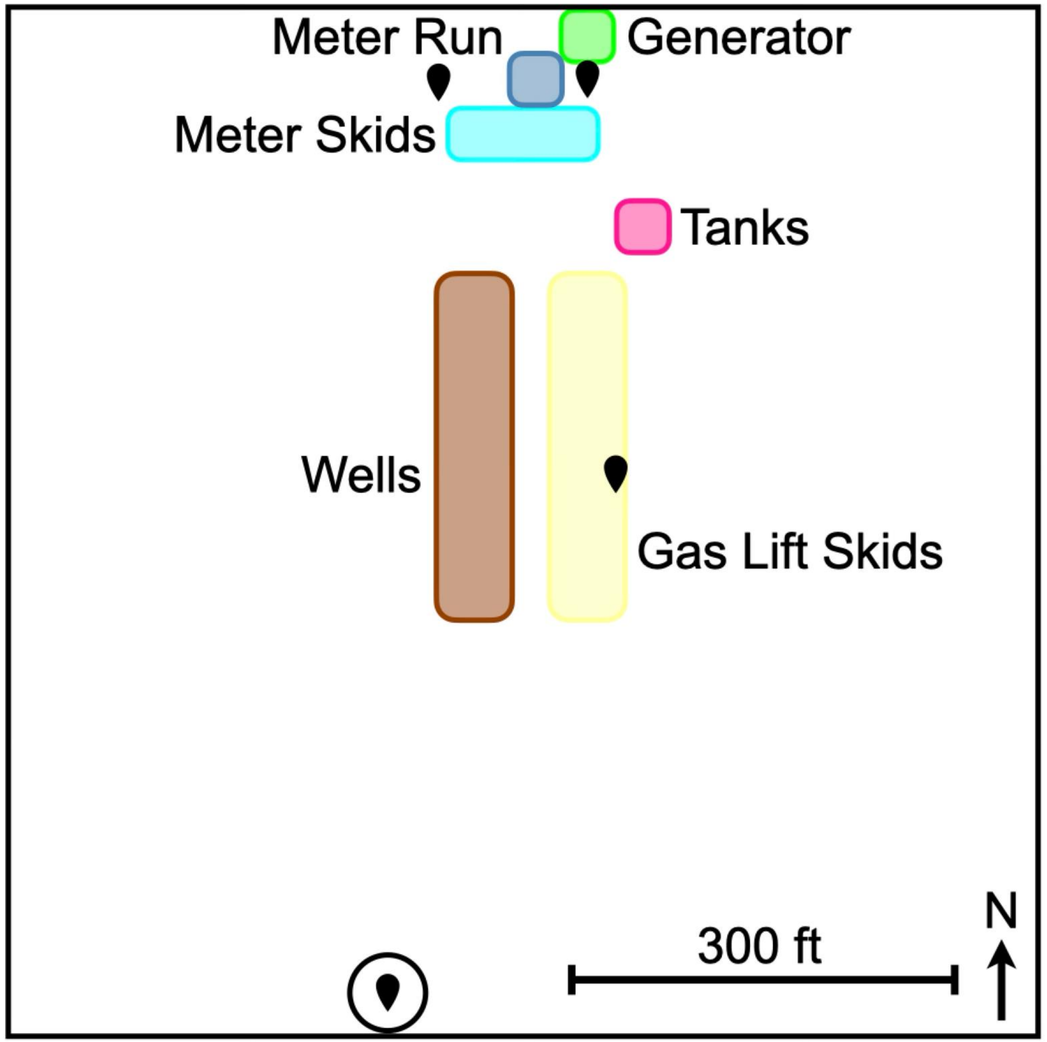


Reconciliation Efforts

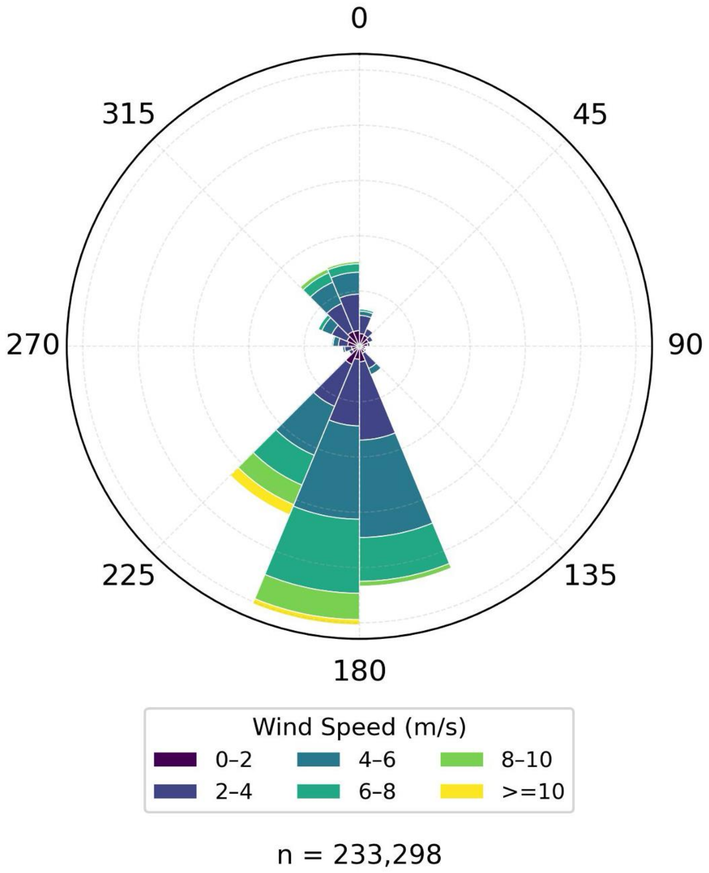
Reconciliation with MAES

- The statistical model and the MAES model report different results, with estimates from the statistical model notably higher than from the MAES model
- Where do these differences come from? How can we reconcile them?
    - One factor: CMS data are not representative of all of Colorado (only 5 sites)
    - Controlling for that, are results still different?
        - We produce MAES and CMS estimates for 5 specific sites and compare distributions

Our Sites: Site 1 Layout

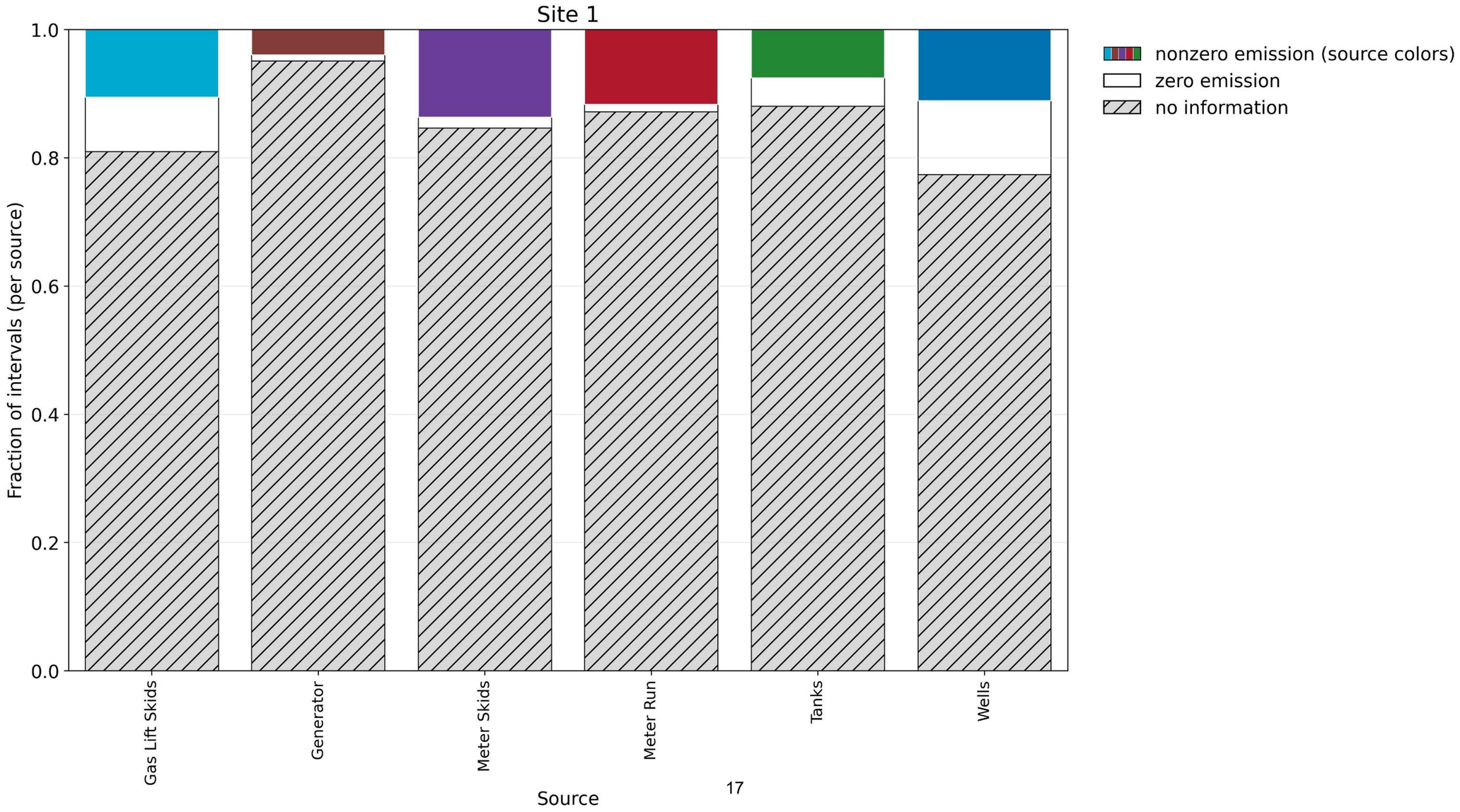


- Anemometer
- CMS Sensor
- Combustors
- Compressor
- Coolers
- Gas Lift Skids
- Generator
- Launcher/Recei
- Meter Run
- Meter Skids
- Pig Launcher
- Separators
- Tanks
- Wells



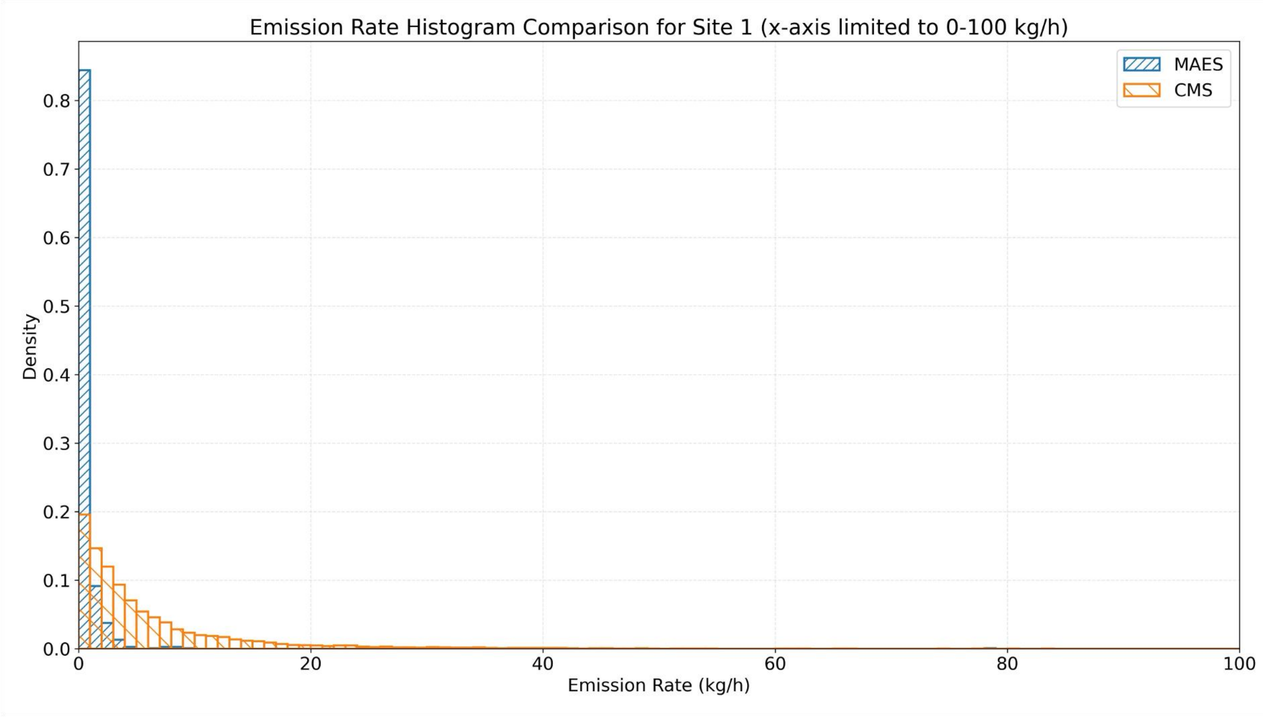
n = 233,298

Our Sites: Site 1 Rate Availability

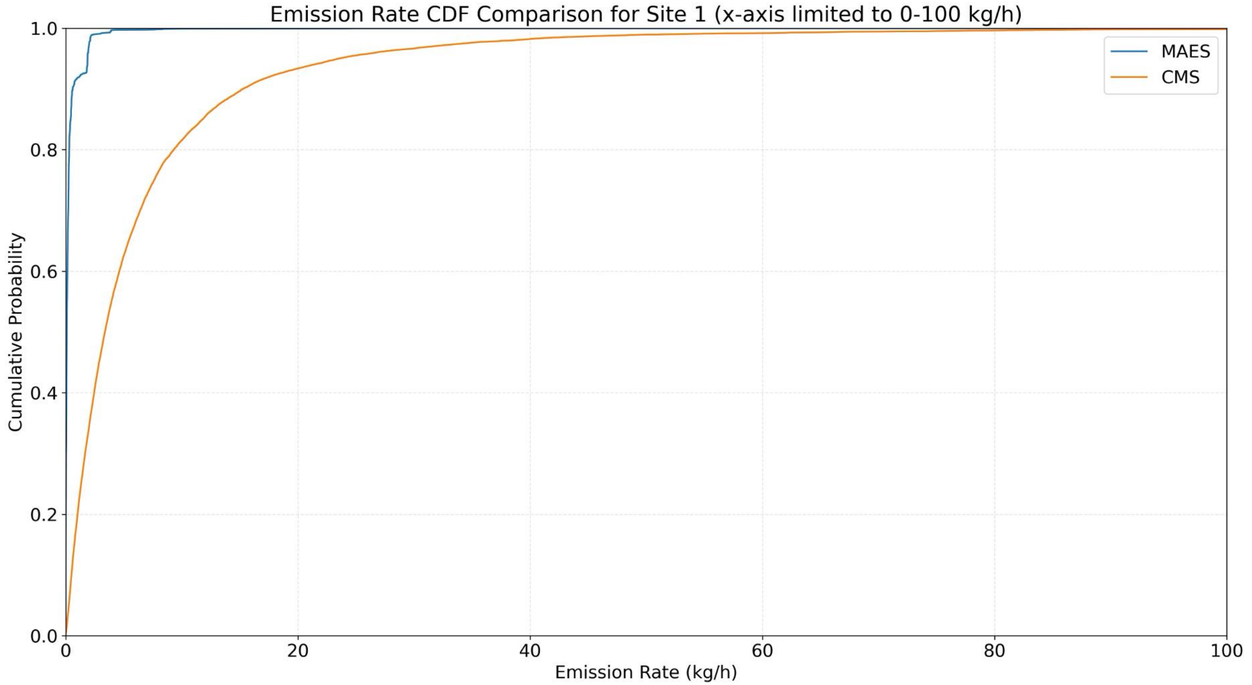


Our Sites: Site 1 MAES Comparison

Histograms

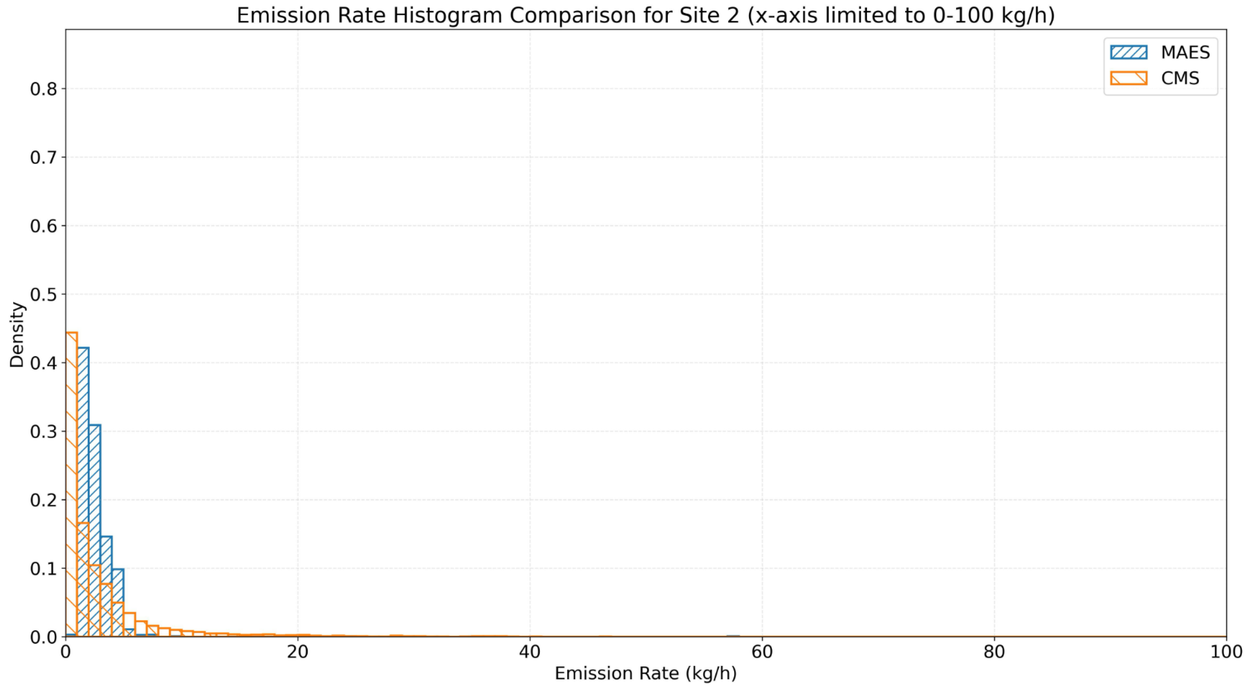


CDFs

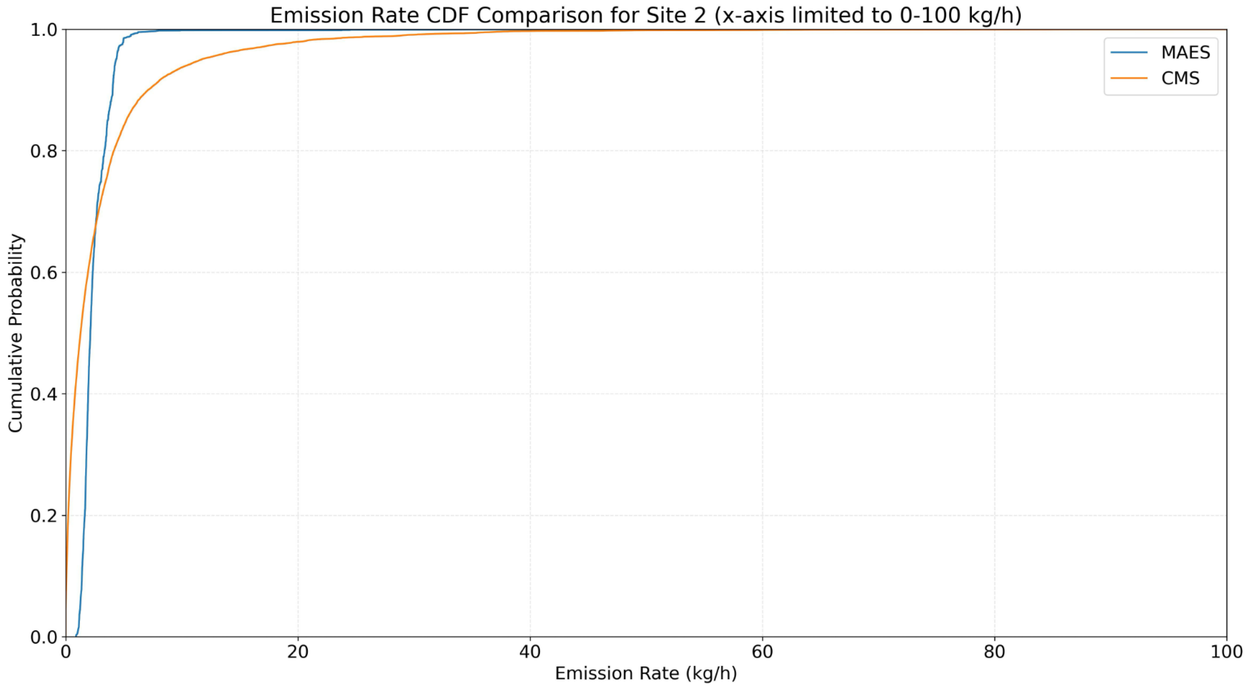


Our Sites: Site 2 MAES Comparison

Histograms

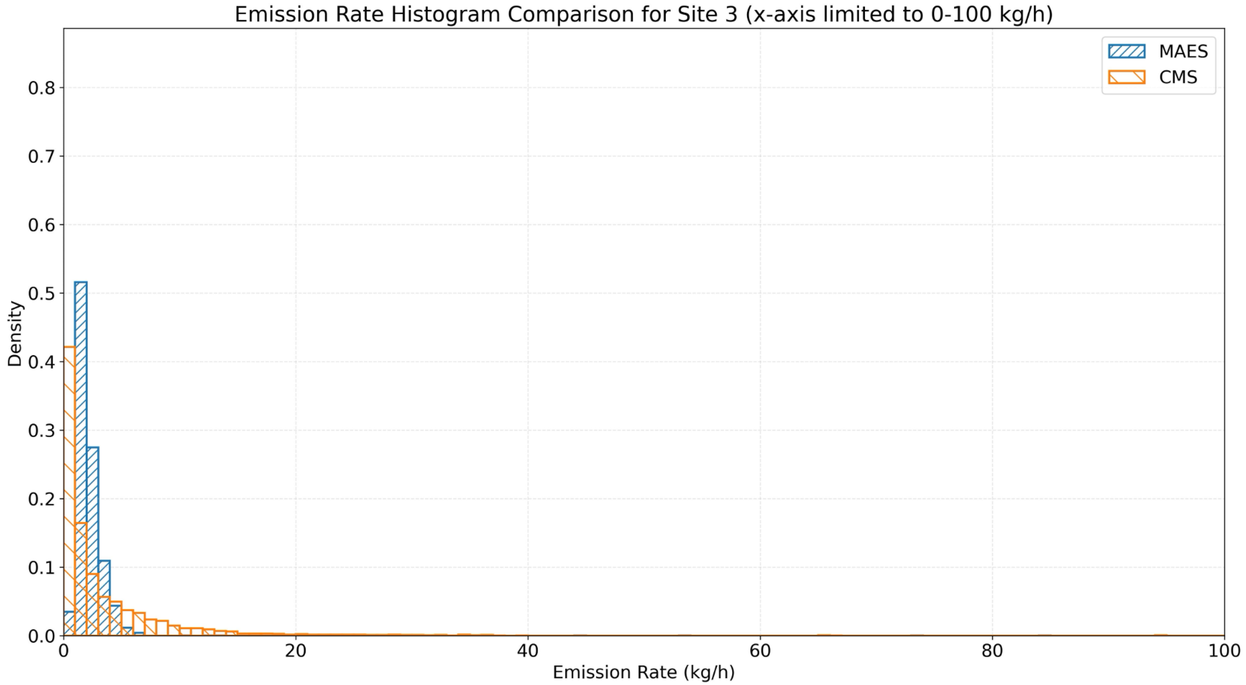


CDFs

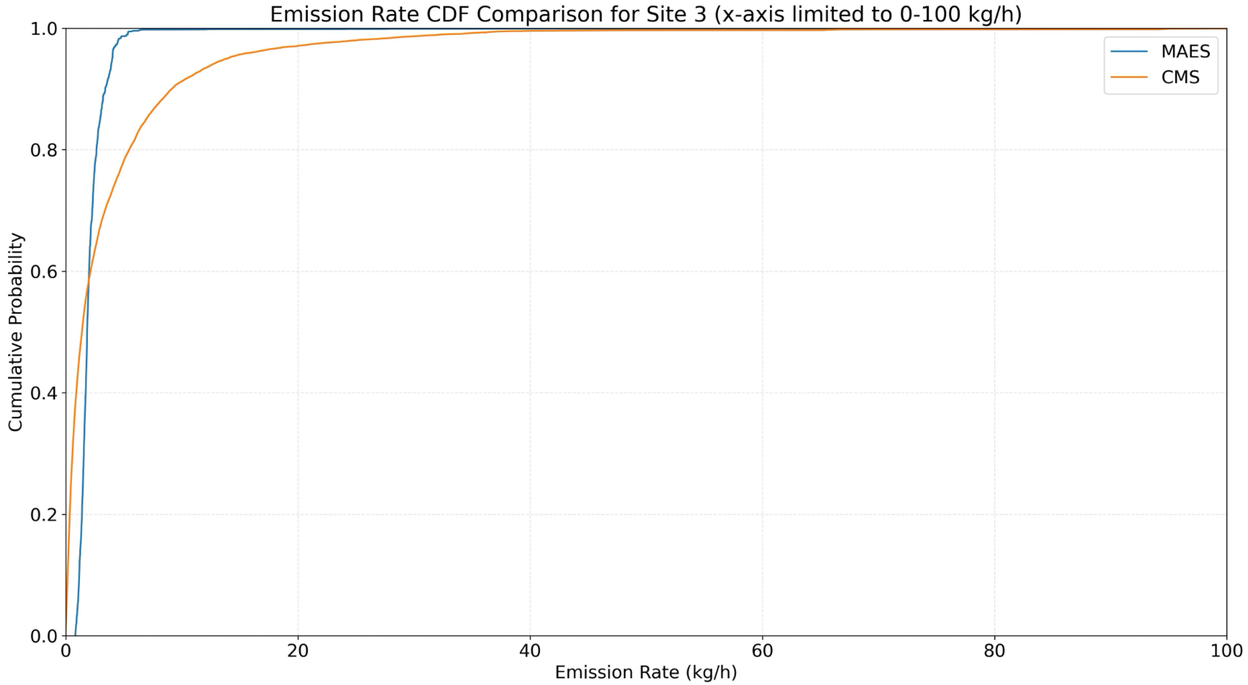


Our Sites: Site 3 MAES Comparison

Histograms

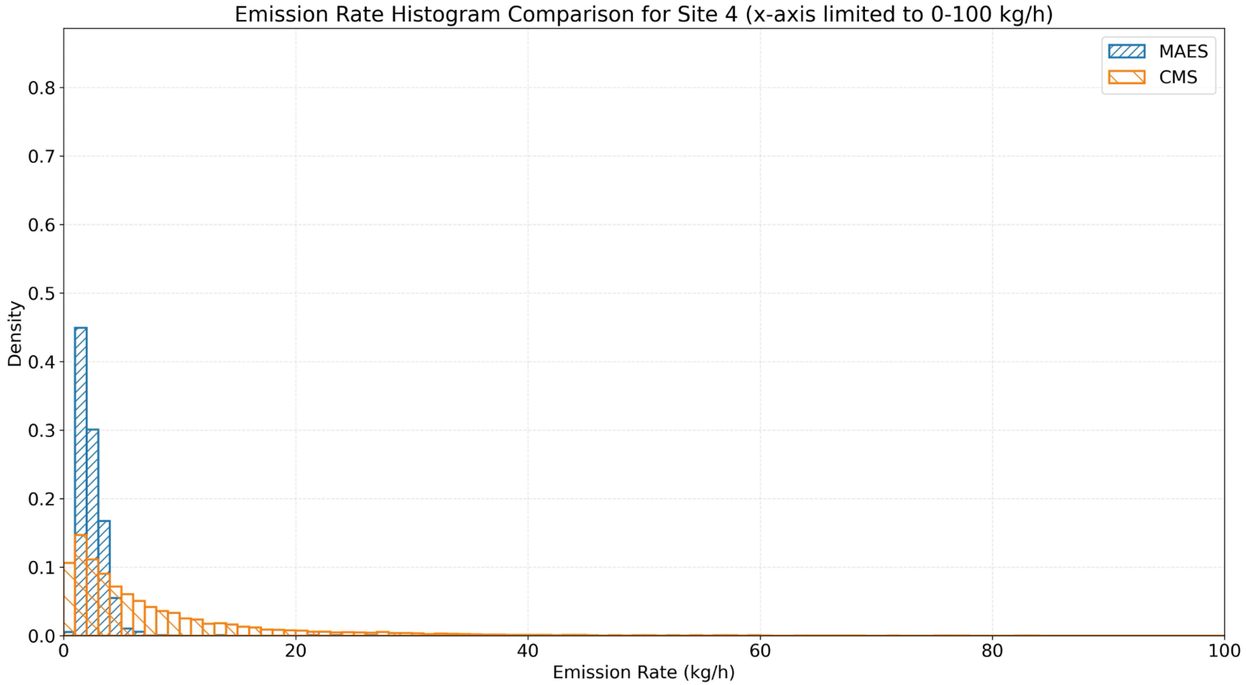


CDFs

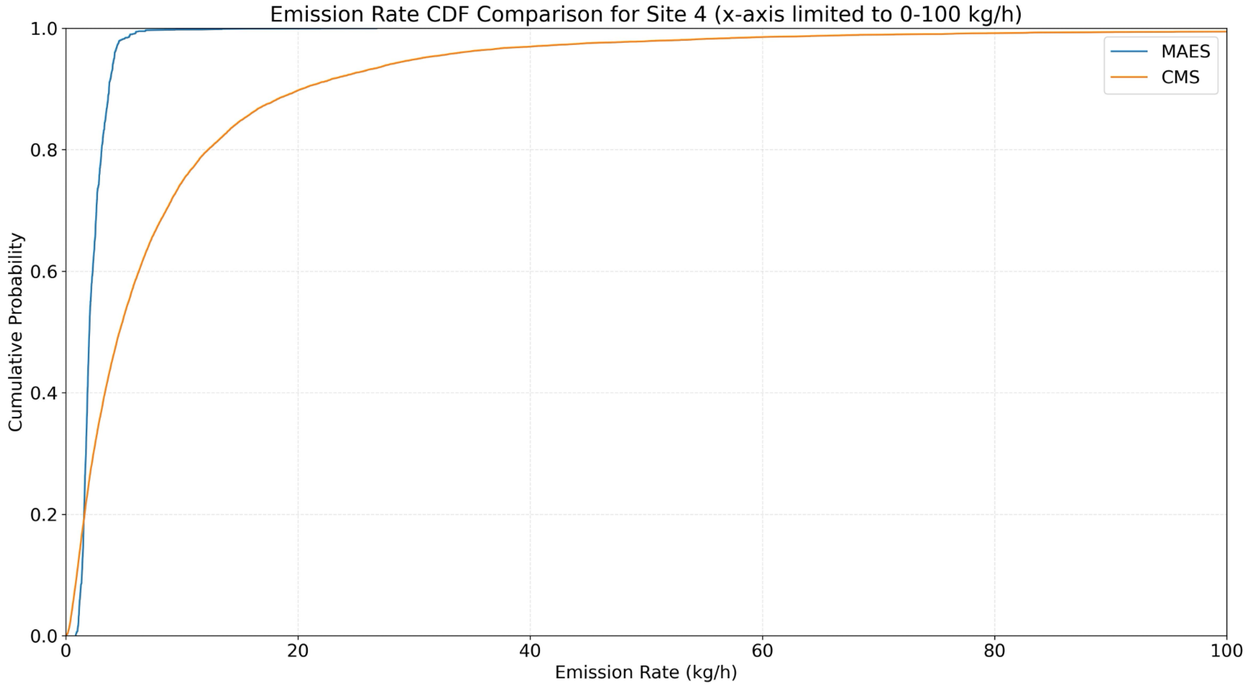


Our Sites: Site 4 MAES Comparison

Histograms

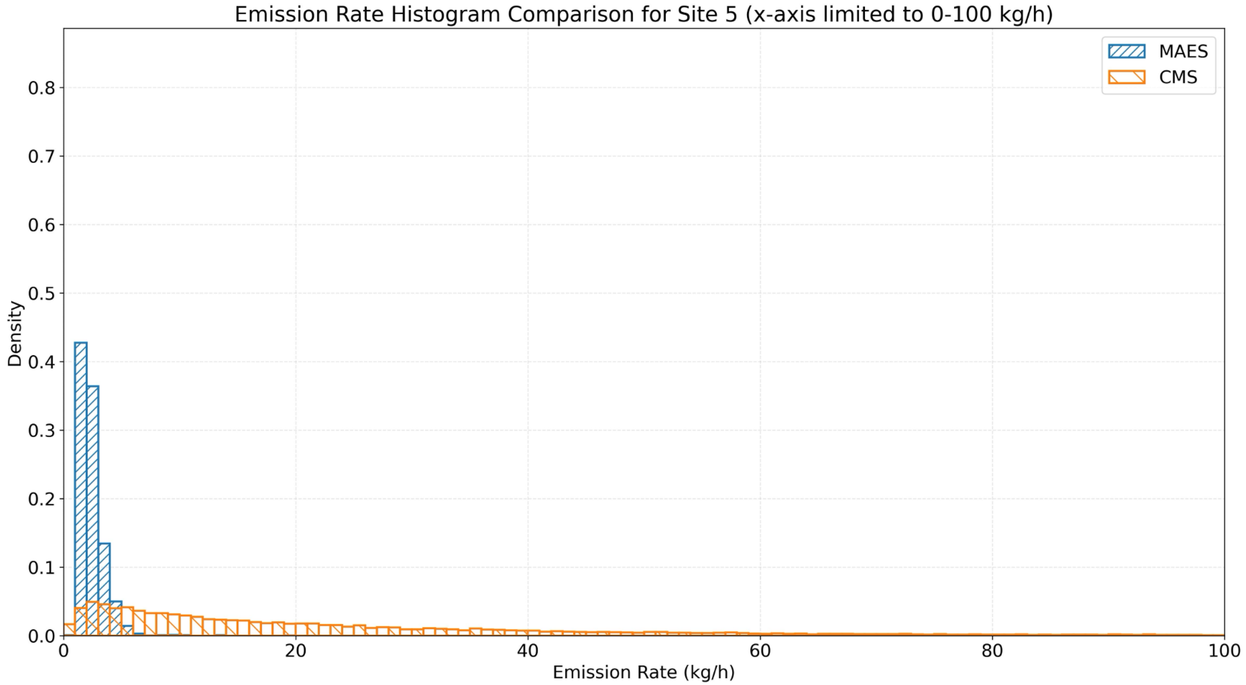


CDFs

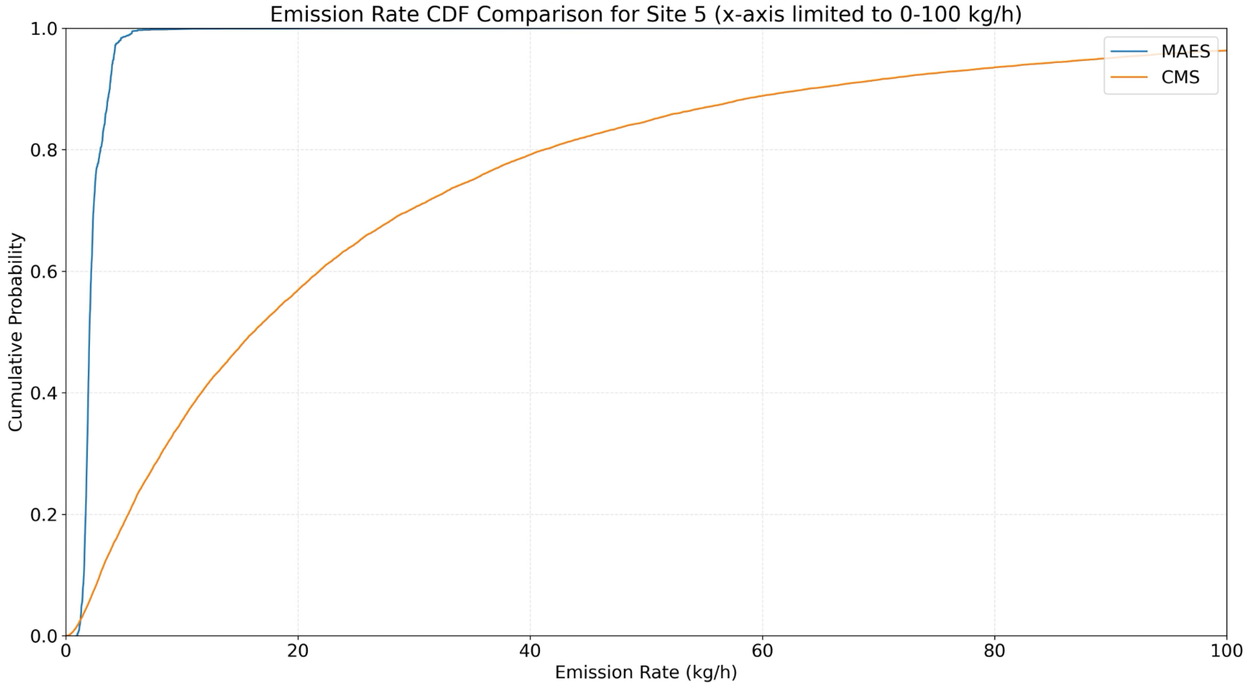


Our Sites: Site 5 MAES Comparison

Histograms



CDFs



Next Steps

- Refine MAES comparison to the equipment-level
- Collect CMS data on more sites to:
  - a) Improve representativeness of CMS data
  - b) Create CMS-informed rate distributions for different subsets of Colorado, and
  - c) Perform more robust comparisons to MAES
- Use statistical methods to segment production sites into groups with similar emissions characteristics and build separate distributions for each

Thank you! Questions?

drichardsdinger@mines.edu