What have we learned from controlled-release studies?

FluxLab Blog 2026 | Post 7



At FluxLab, our experience in controlled-release studies comes from operating the Simulation Facility for Large-Scale Emission Experiments-Landfill (SIMFLEX-LF), a testing facility we built in 2023 on a closed landfill in Petrolia, Ontario. We release methane at known rates from a network of pipes, and measurement teams work “blind”, trying to find and quantify the “unknown” amount. Working with Florida State University and with support from the Environmental Research & Education Foundation (EREF) and Natural Resources Canada (NRCan), we have tested trucks, drones, fixed sensor networks, walking surveys, aircraft and satellites, incorporating up to 21 of them in a single campaign. After each campaign, we publish a report showing performance metrics across methods. The numbers matter, but they do not tell the whole story.

Teams use every source of information they can

Technology developers have a vested interest in performing well in a controlled-release study. Competitors, operators, regulators and investors will be interested in the results, so teams use every legitimate source of information available to inform their estimates.

Early in our landfill testing, we noticed drone teams photographing the site continuously and later using those images to identify pipes, wells and other infrastructure that might indicate the origin of a plume. Initially, this felt like an unfair advantage. It wasn't. If you have a plume of methane and need to determine where it came from, additional evidence is useful, and those teams would use the same information at a real site, often complemented by range finders.

Wind is similar. Global meteorological products cannot always capture the complicated winds around a real facility. When a satellite first detected our releases in 2025, its initial estimates came in as much as 150% too high. Reprocessed with our on-site wind data, they landed close to the true release rate. The satellite detected our releases only three times, so this is an early sign rather than a conclusion, but it shows how much of a "satellite measurement" can depend on data collected on the ground.

This raises an important question: **what exactly is the measurement method?** Usually it is some combination of:

- the methane sensor;
- the wind data;
- site imagery;
- prior knowledge about infrastructure; and
- the algorithm that combines all of these into an emission estimate.

Part of our job is to tease those pieces apart to better understand their relative importance. One drone team that relied heavily on camera footage found every source we released in 2024, without false positives. To test their reliance on visual imagery, the next year we roughly quadrupled the number of visible pipes and decoy features on the ground. The team's sensor and approach were unchanged, but in the second year they correctly located only about six in ten sources. In a follow-up conversation, the team confirmed that visual imagery was an important driver of sensitivity in their process. The point is not to penalize teams for using auxiliary information. It is to identify which

combinations actually make measurements more reliable and then formalize them as part of better measurement practice.

The same instrument can give very different answers

It is natural to assume that a better sensor produces a better measurement. Our results suggest that how an instrument is deployed and how its data are interpreted often matter more.

In 2023, one drone team, using an early path-integrating tunable diode laser absorption spectroscopy (TDLAS) sensor with a small class 1 laser, flew high with widely spaced lines, trolling for readings above a fixed threshold (the way a walking survey would). It missed half the sources and only saw the very largest ones. Another team later used the same low-power sensor but flew low and slow, reading the shape of the plume instead of threshold exceedances. Their performance was one of the best we've ever seen, because of their approach to data collection and analysis.

Continuous sensor networks followed a similar path. In 2023, only one of several networks produced reasonable estimates. By 2025, one network's estimates were centred on the true release rate across 82 separate experiments, using essentially the same sensors in similar locations as the year before. The improvement came almost entirely from better algorithms.

Results from any single campaign are just a snapshot at one stage of method development, not a permanent verdict. Because work practices and algorithms can be updated far faster than hardware, technology performance can improve quickly!

The best-known technology isn't always the best-performing one

Methane measurement is increasingly a competitive business, and some companies are very good at explaining why their approach is the right one before systematic testing has even begun. Sometimes controlled-release studies identify sleeper technologies: the family station wagon that turns out to have a Formula One engine under the hood.

When we began landfill testing, there was relatively little interest from landfill operators in Light Detection and Ranging (LiDAR) or continuous sensor networks. Both quickly demonstrated important strengths. Some continuous networks quantified emissions extremely well, while LiDAR systems showed remarkable sensitivity for finding small leaks.

Neither technology was developed primarily for landfills. Both came largely from oil and gas, and some continuous monitoring teams quickly adapted their algorithms to much larger landfill environments.

That is one reason independent testing matters, and why it helps to have measurement specialists curating the program. If we evaluate only the technologies marketed most aggressively to a particular industry, we will miss promising approaches developed elsewhere.

There probably isn't one winning technology

Nobody would expect a landfill to operate using just one type of machine. Loaders, excavators and dozers exist because they are good at different jobs. Methane measurement is no different.

Our results bear this out:

- The most accurate whole-site methods tell you how much a site emits, but not where.
- Fixed sensor networks estimate emissions around the clock, yet in 2024 they could pinpoint sources precisely only a quarter to half of the time.
- Drones have become our strongest tools for finding individual sources, but they need flyable weather and give a snapshot rather than a continuous record.

So, the question is rarely, "Which technology wins?" A better question is, **"What information do I need, and which tool provides it most efficiently?"** The answer depends on an operator's objectives, budget and regulatory requirements, and it can change over time.

Testing the incumbents can be particularly revealing

Some of the most interesting results come from methods that have been used for years. Two established landfill methods sit at opposite ends of the spectrum.

Tracer correlation has a long history in landfill research. In our experiments at SIMFLEX-LF, it has been among the strongest quantification methods we have tested. We also found that it's possible to break established rules of tracer correlation; it's more flexible than we thought.

The surface walking survey is an important regulatory incumbent, but not a method developed with much scientific evidence that we can find. In 2024, walking surveys using the regulatory 500 ppm threshold found none of the sources we released, while the two drone teams found 88% and 100% of the same sources. Measured by the smallest leak each method can reliably find, walking surveys trail the best drones by roughly tenfold, and aerial LiDAR by more. Tighter spacing and lower thresholds narrow the performance gap, but at the expense of effort and cost. Walking surveys are near the bottom of the list when it comes to performance.

Both methods are familiar and accepted. But how did we come to tolerate such opposites in terms of performance? This situation reveals our human tendency to trust the familiar. These methods have history on their side. They've become waste sector traditions, so we're biased to believe that walking surveys will probably find as many leaks as anything else. In the face of these tendencies, the objectivity of controlled-release tests can be jarring. Familiarity should never prevent us from asking whether substantially better tools have become available.

From waste to coal

These lessons are particularly relevant as we expand our controlled-release work into the coal sector. Coal and waste are very different industries, and open-cut mines present a new set of measurement challenges. But both are traditional industries that have long relied on estimated emissions and are now moving toward direct measurement. We should expect some of the same questions to arise:

- Which technologies from other sectors can be transferred successfully to the coal sector?
- What auxiliary information do methods need to perform well in coal mine environments?
- How much of their performance is related to the hardware, versus other aspects of the method?
- Are there promising approaches that have not yet attracted attention from coal mine operators?
- What combination of methods provides the information we actually need for improved reporting?

Controlled-release experiments are about much more than scoring answers against a known release. We're looking broadly, building a measurement toolbox that helps industries manage methane better. In a future post, we'll ask our collaborators at Stanford, who have been doing this longer than we have, to sketch out some of their observations too.