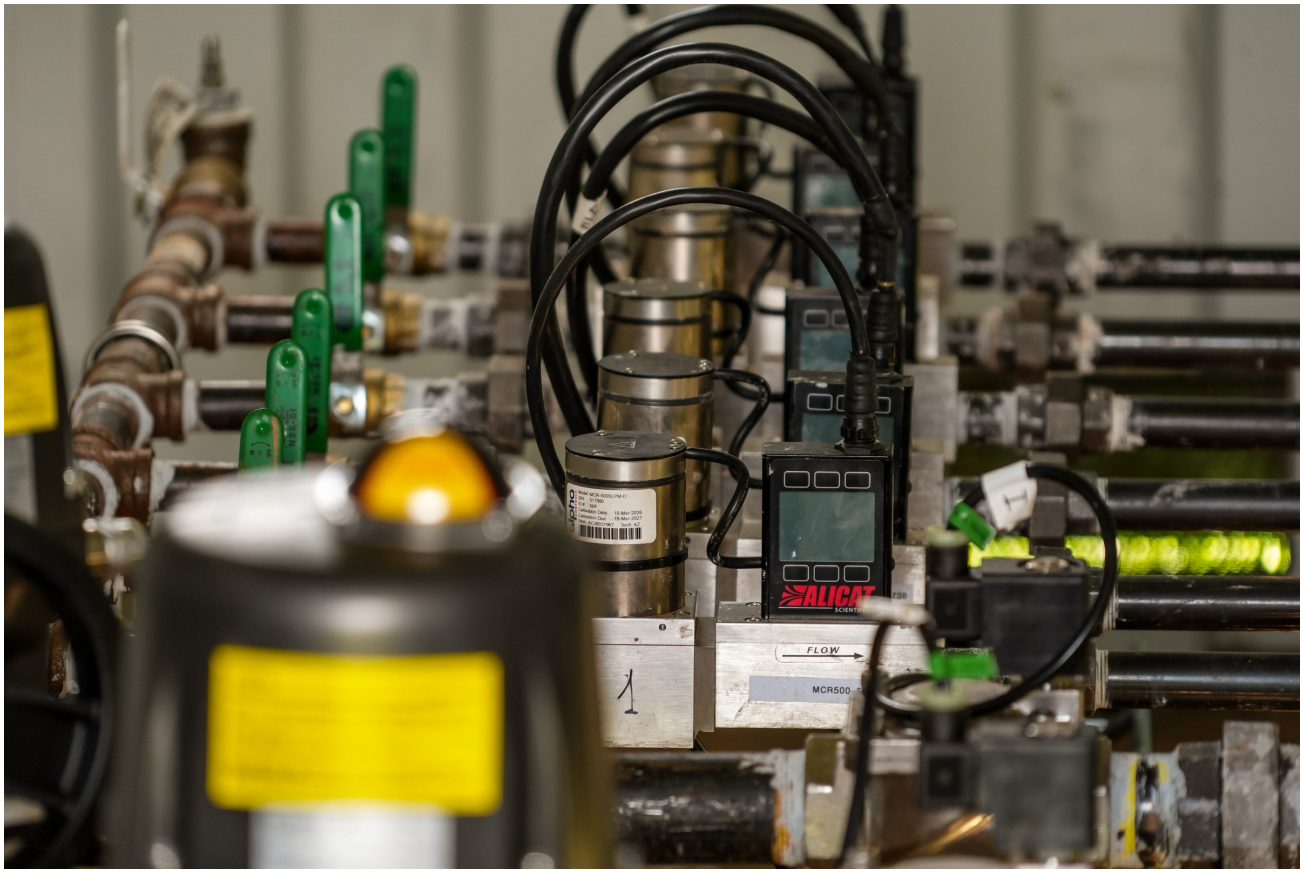
What is the difference between an experimental protocol and an experimental plan?

FluxLab Blog 2026 | Post 3



In our first two posts, we explained why we use controlled releases to test methane measurement at an open-cut coal mine, and which measurement methodologies we hope to evaluate. In short, we release methane at known rates from known locations, and participating teams try to detect and quantify it without knowing the answers.

The next question is how a study like this is organized. Results are only credible if every team knows the rules in advance, and if the releases are designed to test what actually matters. Those two needs are met by two very different documents: the **experimental protocol** and the **experimental plan**.

The distinction matters for participating measurement teams, because the documents serve different purposes. It also matters for anyone interpreting the study later, because one explains how results will be judged, while the other determines what questions are tested.

The simplest way to think about it is this:

- The protocol is the public rulebook.
- The experimental plan is the confidential playbook.

The Experimental Protocol

The protocol defines the formal rules of the study. It is, in effect, an agreement between the team operating the controlled-release site and the participating measurement teams. Before anyone flies an aircraft or drone, schedules a satellite observation, drives a measurement truck downwind, or assembles a sensor network, they need to understand how the study will work and how they will be evaluated.

The protocol explains things such as:

- what participants are expected to report;
- which performance metrics will be evaluated;
- whether source location, emission rate, or both must be submitted;
- the required data formats and submission deadlines;
- how measurement intervals are defined;
- what information remains blind; and
- how results will be scored and reported.

Controlled-release studies have evolved substantially over the past decade. Early experiments, including important oil and gas studies led by Stanford, often focused on relatively simple questions: measurement bias, precision, or minimum detection limits. Those concepts are now treated with greater nuance. In our first post, we introduced several of the metrics now in common use, including the 90% probability-of-detection emission rate and the classification of detections as true positives, false positives and false negatives. Quantification performance is likewise evaluated in several ways, from bias relative to a one-to-one relationship with known emissions, to precision across repeated estimates, to the spread of individual errors across all releases.

The protocol specifies which of these outcomes matter for a particular study. That is important because many measurement systems can do more than one thing. A methodology may detect, localize and quantify methane, but the study may designate one of those functions as the primary

performance claim. Participants need to know what they are being asked to optimize and what will ultimately be evaluated.

Where do the protocols come from?

The most advanced methane controlled-release protocols are those that have emerged from the oil and gas sector, where measurement methodologies increasingly need to demonstrate performance under realistic operating conditions.

That means representing more than an isolated leak from a single pipe. Real facilities may have several active sources, variable background emissions and routine venting. A methodology may therefore need to quantify an incremental emission above a non-zero baseline rather than simply distinguish between “methane” and “no methane.”

The METEC facility at Colorado State University has played a leading role in this development. Its protocols support rigorous testing of methane measurement systems and may form part of the evidence used in regulatory approval processes. They have been developed and reviewed with input from universities, measurement companies, non-governmental organizations and other technical experts. Stanford has adapted elements of this framework in more recent controlled-release work, as has StFX for our landfill campaigns.

For our Australian study, we adapted the protocol from those landfill campaigns. The changes were mostly reflecting the layout and geometry of an open-cut mine, while the core rules stayed the same. The protocol will be made available well before testing so that all participants understand the rules. Once the campaign begins, those rules cannot be changed simply because the results are surprising or because one type of methodology performs better than expected. That stability is central to fairness and scientific credibility, and it is why the protocol should not depend on the type of site. What changes in a coal mine is the experimental plan.

The Experimental Plan: What We Actually Do

The experimental plan is very different. It contains the detailed sequence of releases that we intend to run, including which nodes should be active, at what rates, for how long, and in what combinations. In other words, it contains the answers.

For that reason, the experimental plan is confidential and is normally known only to a small group responsible for designing and operating the experiments.

Where the protocol is fixed and public, the experimental plan is restricted and dynamic.

How is the plan developed?

The planning process begins by identifying the scientific questions and the types of measurement methodologies we want to evaluate. Participating teams are then recruited and asked to complete a detailed technical questionnaire. We need to understand not only what each system is designed to do, but how it actually operates. The questionnaire may ask for:

- the expected measurement duration;
- minimum and maximum measurable emission rates;
- expected detection thresholds;
- wind-speed or weather limitations;
- spatial coverage;
- required flight or driving patterns;
- deployment and setup time;
- whether the system can measure continuously;
- whether repeat estimates can be submitted during one release; and
- safety or operational constraints.

This information is essential. A satellite may take a nearly instantaneous image, while an aircraft survey may require an hour and a ground network may operate continuously. A drone may need several battery changes. A truck may require access to a suitable road at the correct downwind location.

At an open-cut mine, the plan also depends on how air moves through the pit. We are using wind modeling to help decide where to place release nodes and receptors. It also helps us set go/no-go criteria: the weather conditions under which measurements from inside the pit can be trusted.

We need to create release events that are long enough for the required methodologies to complete their measurements, but not unnecessarily long. In a pit, releases must also run long enough for methane to reach steady state, where the amount added equals the amount leaving. We also need enough repetitions, across an appropriate range of emission rates and environmental conditions, to support meaningful statistics. The result is an experimental matrix.

Building a common experimental matrix

Many methodologies can measure the same release simultaneously. During one experiment, we might have:

- trucks working on downwind roads;
- one aircraft above the site;
- one or more drones, where airspace and safety allow;
- a satellite passing overhead;
- continuous sensors operating around the site; and
- open-path or tower systems recording throughout.

This can make controlled releases remarkably efficient. A single known release may generate useful data for many teams. There are practical limits, however: some methodologies, such as drones and aircraft, require exclusive access to particular areas for safety.

The release duration is often set by the slowest high-priority methodology. Faster systems may be asked to submit multiple results during the same constant-release interval. This helps us examine repeatability without changing the controlled conditions.

The matrix also needs to create comparable opportunities among teams. If several aircraft systems visit at different times, we may need to repeat a similar range of emission rates and source configurations for each one. Continuous systems, by contrast, may observe almost every release and accumulate many more replicates across changing weather. That difference is not necessarily a problem, but it needs to be understood when results are analyzed.

Designing experiments that test real vulnerabilities

A good experiment does not simply operate a release system at random rates. It begins with a scientific hypothesis. We need to understand how each methodology works, what assumptions it makes, and where it may be vulnerable. For example, a drone team searching for active sources may use ordinary visual imagery as well as methane measurements. If the only visible pipes or release points are the active ones, the visual information may make the localization task unrealistically easy. We can make the test more rigorous by installing numerous inactive decoy nodes. The participant must then rely more heavily on the methane measurement rather than simply identifying the obvious infrastructure.

Tracer-correlation methods provide another example. They work best when the tracer release is spatially well matched to the methane source. In a real industrial setting, the true methane distribution may not be known. A team might place the tracer near the centre of the site even though most emissions are coming from one side. We can reproduce that problem by creating a strongly asymmetric methane release. The purpose is not to trick participants for entertainment. It is to test a known methodological vulnerability under conditions that could occur in practice.

Open-cut mines offer a third example. Most methods assume that methane moves steadily from the source to the sensor. A pit does not always behave that way. Our simulations suggest that methane can pool in the pit overnight and then escape fairly quickly once the ground warms in the morning. A team measuring during that morning release could mistake hours of stored methane for a higher emission rate. By running releases through these transitions, we can see whether a methodology recognizes when this is happening.

This is one of the main benefits of controlled-release testing: we can create difficult but interpretable situations and determine where a methodology begins to fail.

Why the plan is developed late, and then keeps changing

Although experimental themes are developed from the beginning of the project, the detailed plan normally comes together quite late, often within the final weeks before a campaign.

By then, we have a confirmed participant list, travel schedules, insurance approvals, aircraft availability, satellite opportunities and a much clearer understanding of which systems will actually be present.

Even then, the plan remains fluid. Controlled-release campaigns are dynamic. We may have many measurement teams operating at once, and every change in flow rate or source location must be communicated and logged precisely. At the same time, aircraft schedules shift, equipment fails, winds change direction, and weather conditions become unsuitable. A carefully designed release scheduled for the afternoon may become impossible because the wind turns. A satellite observation may be lost to cloud. An aircraft may arrive a day late. A drone may require a longer operating window than expected. The experimental team therefore has to reorganize releases quickly while preserving the scientific intent.

Adapting the plan this way requires the people managing the experiment to understand both the study hypotheses and the practical operation of the measurement systems. Otherwise, a seemingly harmless schedule change can undermine the comparability or value of an experiment.

How the two documents come together

The protocol and the experimental plan ultimately serve the same study, but they do very different jobs. Participants need to read the protocol carefully because it tells them what the study expects of them. They will not see the experimental plan, because it contains the information they are being asked to measure. They will, however, see its influence in the final report. The hypotheses and experimental themes built into the plan become the structure of the eventual scientific analysis.

A strong controlled-release study depends on both documents: a stable and transparent set of rules, and a flexible experimental design created by people who understand the technologies well enough to test them properly.