https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

# Coverage of mobile methane leak surveys of gas distribution networks: the impact of wind-aware scheduling on the coverage distribution

David Ball[1], Nathan Eichenlaub[1], Karin Ohman[1], and Eric Chan[1]
[1]Project Canary, Denver, Colorado, USA
**Correspondence:** David Ball (david.ball@projectcanary.com)

**Abstract.** Mobile leak surveys of gas distribution networks typically drive each street six times, two passes a night, and often run those nights in close succession. Each pass's coverage (the pipe it is able to survey) depends on the wind direction, which is highly autocorrelated in time. This raises a concern about the standard practice: when passes are packed into a few nearby nights, does the wind direction vary enough for the survey to "see" the entire intended pipe, or do the repeated passes simply re-survey already-covered pipe and miss the rest?

To investigate this, we use 3,992 real survey drives, each paired with a plume model that computes the fraction of the pipeline a drive could survey under a given night's wind. For each drive we run a Monte-Carlo simulation, repeatedly drawing historical wind from other candidate nights and computing the resulting distribution of coverage under two deployment strategies: "wind-blind" surveys that drive on random nights within a week, and "wind-aware" surveys that choose nights for complementary wind directions. We find that six blind passes over three closely spaced nights cover a median of 81% of pipe. The distribution has a low tail: a hard network, or any network caught in an unlucky week, is left with roughly a third of its pipe uncovered, and that tail deepens for the smaller leaks, while line-of-sight blockage (a heuristic but physically motivated treatment of obstructions' impact on plume detectability) lowers coverage further still. Because the wind varies little over a single night, coverage is set by how many distinct nights a survey samples rather than by its pass count. Choosing those nights by the wind, using only conditions available within a one-to-two week window, removes the low tail and covers more with fewer passes: four passes over two chosen nights beat six driven blind, and three chosen nights reach about 93%.

## 1 Introduction

Mobile advanced leak detection (AMLD) has become a standard tool for compliance surveys of gas distribution networks: an instrumented vehicle drives the streets above the buried pipelines while a fast methane analyzer flags enhancements in concentration that indicate a nearby leak (von Fischer et al., 2017; Weller et al., 2018). In practice a street is not driven once but several times. A vendor application brief specifies "multiple (typically six) passes on each street . . . on at least two different nights" (Picarro, Inc., 2019); in conversations with distribution operators we have found the standard practice is to employ two passes a night, one for each side of the street, spread across two or three nights within about a week. That cadence is generally consistent across the industry. Driving the same stretch of road several times is meant to maximize coverage (the fraction of

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

pipe over which the survey could actually have detected a leak), and give the AMLD platform multiple chances to catch a leak in a given location; recognizing the fact that detecting the plume from a leak on a given pass is not guaranteed, and may have some element of randomness. Coverage, rather than the pass count itself, is what a survey is ultimately judged on, and it is already what operators document, with coverage maps and hand-surveys of the remaining gaps. We frame our analysis in those terms throughout.

Where does the six-pass default come from? Two independent analyses of the problem reached similar conclusions using different lines of argument and underlying data. The first is the vendor protocol itself: drive six passes over at least two nights to maximize wind direction variability, then map the resulting coverage and hand-survey whatever gaps remain (Picarro, Inc., 2019). This is not naive: it identifies the right variable (wind) and validates against the right quantity (coverage). The second comes from the literature. Luetschwager et al. (2021) estimate that five to eight drive-bys capture more than 90% of leaks, a figure they derive by treating each drive-by as an independent Bernoulli trial, in effect an independent coin-flip at catching the leak, and accumulating detection probability across passes. These two analyses converge to roughly the same conclusion while measuring different things, pipe covered in one case and leaks captured in the other, which is part of why the convention feels so firmly settled. The independence assumption in particular has gone largely untested; later work repeats the five-to-eight figure without re-examining it (e.g., Maazallahi et al., 2023, who restate it as a per-location detection probability while noting that mobile detection depends on wind direction), and to our knowledge no subsequent work has examined the assumption directly.

There is a physical reason to be uneasy about treating each pass as independent. Whether a given length of pipe is detectable on a given pass is set largely by the geometry between the wind and the road: the vehicle detects a leak only when it drives through the leak's plume, and where that plume goes is set by the wind. This wind-conditioned detectable region is not a new idea; it is the subject of a patent family (Tan and Rella, 2017, 2025), which computes it for a single pass, but the patents do not analyze how coverage accumulates across passes under varying conditions. What decides whether a repeated pass adds new coverage or simply (and less helpfully) repeats the last one is how far apart in time the passes fall. Because wind direction is highly autocorrelated, it varies little over the course of a single night, so repeated passes of the same street on the same night largely revisit the pipe the earlier passes already covered rather than reaching new pipe. Repeating a pass can still be worthwhile: if a single pass downwind of a "covered" segment has some probability of "missing" a detection (e.g., a particularly low-concentration instantaneous realization of the turbulent plume structure, or if the plume is disrupted by a nearby vehicle passing by) or because the leak vents intermittently and is quiet as the vehicle goes by, a second pass gives another chance at the same pipe. That benefit is real but limited: a repeat pass can only help on the regions of pipe that are "covered" by the geometry of the drive, and does nothing for pipe that is downwind of the survey route. Repetition improves detection on top of coverage, but it cannot make up for gaps in coverage of pipe the wind never carries a plume from; we return to this imperfect per-pass detection in §3.4.

The controlled-release literature, the field's gold standard for understanding detection probability, cannot see this effect: controlled-release campaigns place their measurement transects downwind of the source or retain only the downwind passes while analyzing the data, so no published probability of detection for street-level surveys is conditioned on wind direction

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

relative to the road (Tian et al., 2022; Venkata Rao et al., 2026; Gundapuneni et al., 2026). In other words, they answer the different question of: "provided that the vehicle drove downwind of an emitter, what is the probability that the source was detected?".

The authors of the independence assumption identified this gap themselves: lacking reliable wind measurements, Luetschwager et al. (2021) could not distinguish upwind from downwind sampling attempts, and noted that their detection rates were therefore a lower bound, and anticipated that surveys incorporating wind "could quantify the spatial coverage of surveys." In this study, we aim to address exactly that. The closest existing work to put wind inside the routing loop is Gao et al. (2022), who plan survey routes from a wind forecast by drawing a straight downwind ray from each known point source (a well or facility) to the nearest road. That approach looks ahead from a forecast, treats discrete point sources, and is purely geometric, with no dispersion model and no treatment of how many passes are needed. Our approach differs on each of these points: we treat a continuous pipe network rather than discrete sources, decide detectability from a dispersion model rather than a straight-line ray, and work backward, drawing from historical wind data, to simulate what direction-blind surveys may leave uncovered.

This study utilizes a dataset containing 3,992 survey drives, each with wind measured on the moving vehicle alongside pipeline geometries associated with the driving regions. To explain why direction-blind repetition can result in suboptimal coverage, we use these drives to resolve two distinct timescales of wind direction variability, compute the wind-conditioned coverable set on real distribution networks, and compare direction-blind against wind-aware surveying at scale. Earlier studies either lacked the wind measurements needed to resolve the effect (Luetschwager et al., 2021) or used vehicle-measured wind only to check predicted plume intercepts at discrete oil and gas sites rather than to assess coverage of a continuous pipe network (Gao et al., 2022); these drives are what let us do this after-the-fact, network-scale analysis. The coverage numbers we report are simulated rather than observed: each drive contributes its route and network geometry once, and we replay that route under winds drawn for other candidate nights to ask what a given scheduling rule would have covered. The results here characterize those scheduling rules applied to real routes and regional winds, not the coverage any of these surveys actually achieved as they were driven.

This paper makes four contributions: first, as a check on the premise behind pass counting, we test the independence assumption directly in the data underlying it, and find the passes are strongly correlated rather than independent (Section 3.1). Second, using our own survey drives we identify the mechanism: two timescales of wind spread, tight within a night and widening only across a span of days (Section 3.2). Third, we compute coverage at a fixed pass count as a distribution (via a Monte-Carlo scheme), compare direction-blind driving against a wind-aware upper bound, and show the comparison is robust across a grid of sensitivity, line-of-sight (LOS), and residual-detection assumptions (Sections 3.3–3.4). Fourth, we ask how far ahead one must be willing to choose survey nights, and find that a one-to-two week window captures nearly all of the available benefit for the particular weather patterns in the regions this study's data comes from (Section 3.5).

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

## 2 Data and methods

### 2.1 Survey drives

We work with 3,992 processed survey drives, each roughly an hour of 1 Hz records with on-vehicle anemometry corrected to the ground frame, drawn predominantly from a single mid-latitude US territory in the Midwest (about four in five drives), with the rest from other US regions and a small number from other climates that we return to in the regional comparison of §4. This pool is large enough to characterize how coverage varies from site to site, and the mechanism, coverage-distribution, pass-count, and scheduling-window analyses (§3.2–§3.5) all run on it. The pass-correlation test (§3.1) is performed on a different dataset: the external 12-city drive-by records of Luetschwager et al. (2021) rather than on our drives.

### 2.2 Two denominators for coverage

Reporting coverage as a fraction of pipe first requires deciding which pipe counts as a drive's target, and there are two natural choices. The first uses the planned route's assigned pipe; for some drives the utility records the pipe the crew was supposed to survey that night, and we take that recorded route's pipe set as the denominator. The second uses the drive itself, using all the pipe within 100 m of the vehicle track, so the target is the pipe the vehicle actually passed near rather than the pipe it was assigned. In principle, the planned routes' pipe sets are the cleaner target, however this information is only filled out for a small subset (a few hundred out of several thousand) of the total drives. Because of this, we opt to simply use the buffered drive approach to define the coverage denominator (which can be computed for any drive and does not require specific data entries that may be missing). We settle on the 100 meter buffer specifically because it makes the driven-track denominator match the recorded work-order mileage closely (within about 5%) on the drives where both exist. This approach is arguably the more faithful measure of what a night's drive could have found, since it follows where the vehicle went rather than where a plan said it should go. We therefore use the driven-track pipe as the primary denominator for the at-scale comparison, and keep the planned-route pipe as a cross-check on the drives where the route was recorded. The two give the same comparative answer: the wind-aware gain reported below holds on both, and only the absolute coverage levels differ, because they count different pipe.

### 2.3 Wind: measurement and other candidate nights

Each drive has its own measured wind, in the ground frame and filtered, and that is what we use for the as-driven analysis. To ask what other nights would have looked like, we need wind for nights the crew never drove. We take it from ERA5, a global atmospheric reanalysis that reconstructs past hourly wind from weather observations, using its night-hour winds (2022–2024) at the drive's location to ask what other nights would have offered. ERA5 reports speed at 10 m, well above the vehicle intake, and a naive log-law correction over-predicts because it ignores nocturnal stability; we instead rescale the ERA5 night-hour speeds empirically so their median matches the drive's own measured median, leaving direction untouched. It's worth noting that this rescaling will affect the footprint size, but not the directional structure that drives our result. In order to assess how

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

variable wind directions are over a given time period, we use the measured anemometry and consider the spread of wind direction over two distinct timescales: within one night, and across a week. We summarize a set of wind directions by their circular standard deviation: the directional counterpart of an ordinary standard deviation, accounting for angle wrapping, with small values meaning the directions cluster in a narrow arc and larger ones meaning they are spread more widely around the circle (it grows without bound as the directions approach a uniform spread around the compass). The direction is tight over each roughly hour-long drive, with a circular standard deviation of only about 13°, a band narrow enough that every pass points plumes in nearly the same direction; we take this as evidence that the direction is also stable over the span of a single night, from which the drives are drawn. Over a longer span it widens, to ∼77° across a week. We note that these results are specific to the particular weather patterns in the locations that these drives occur and not necessarily representative of wind variability more generally. We return to this point in §4 where we provide a brief analysis of wind variability across distinct geographic regions where our drive data is taken.

### 2.4 Wind-conditioned coverage model

Whether a survey detects a leak on a given street depends on two distinct factors. The first is feasibility: was the leak detectable at all on that pass? Feasibility comes down to the geometry of the vehicle's position relative to the leak and the wind direction, and effectively asks whether the vehicle passed close enough downwind of the leak to detect it. The second is the probability of detection given a feasible encounter: once the vehicle does pass through a plume the geometry says it should see, how often it actually registers the leak. Most controlled releases address the second factor but not the first, because their transects are placed downwind (or their analysis keeps only downwind passes). Our model computes the feasibility term explicitly, from the measured wind and the road geometry, and folds the conditional-detection term into a single per-encounter probability $p_{\mathrm{enc}}$. The coverage we report is the product of the two.

At the center of the analysis is a model of which pipe a drive could have detected a leak on, and we build it by inverting the usual forward picture. Rather than disperse a leak outward from each pipe segment to the vehicle, we place a hypothetical unit source at the vehicle, reverse the wind, and read the concentration it produces at every nearby pipe segment at once; by reciprocity this equals the methane the vehicle would have measured had a unit leak sat at that segment. Each vehicle position is thus the source and each short length of pipe a receptor. A steady-state Gaussian plume model, the standard textbook description of how a continuous release spreads and dilutes downwind, run for stable night air (Pasquill class E, the stable end of the standard atmospheric-stability scale), supplies the concentration for every source–receptor pair under the wind measured at that instant. A receptor is covered if that modeled concentration clears a detection threshold at any timestep on the drive, and the covered set for the drive is the union over all its timesteps. Figure 1 walks through the construction on one drive.

Because the Gaussian plume is linear in emission rate, the detection test depends on the leak rate $Q$ and the instrument's minimum detectable level MDL only through their ratio, so we collapse the two into a single sensitivity axis, $s = Q/\mathrm{MDL}$ (in $\mathrm{g\,h^{-1}}$ per ppm). Low $s$ is a demanding case, a small leak or an insensitive instrument, while high $s$ is an easy one. This keeps the results vendor-neutral: we never have to state an MDL, and a reader with any instrument and any leak size of interest can locate themselves on the $s$-axis. We report at $s = 360$ as a central case: a round mid-range value whose plume footprint reaches

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

about 80 m downwind, sitting between the demanding and easy ends of the range we sweep, $90/360/1440$ (about a $16\times$ range, footprint reach from 25 to 215 m). In general, leaks from distribution pipelines are expected to be small on average, so the operationally dominant cases sit toward the low-$s$ end of the swept range, which is where the advantage of choosing nights that we report is largest. A related single-pass construction, in which the same ratio together with wind speed sets a maximum detection distance, appears in Tan and Rella (2017). The one caveat is that the linear collapse of $Q$ and MDL into $s$ holds only for a single-threshold detection rule; a nonlinear rule (ethane-ratio confirmation, multi-point persistence logic) would break it and must be modeled explicitly.

As noted at the start of this section, the relative geometry of vehicle position, pipeline, and wind direction dictates the feasibility term: is the pipe upwind and sufficiently close to the vehicle that a leak would be detectable? This is distinct from the conditional-detection term, and real surveys do miss leaks on segments that the geometric coverage labels coverable, for reasons the model does not represent: an obstruction, a nearby vehicle crossing the plume, the instrument's duty cycle, or intermittency in the underlying emissions. This is the $p_{\mathrm{enc}}$ introduced above, which we apply as a per-encounter detection probability, a flat multiplier on the geometric coverage. Mathematically, this is expressed as

$$P(\text{detect} \mid \text{pass}) = \begin{cases} p_{\mathrm{enc}} & \text{if the forward model marks the receptor covered on this pass,} \\ 0 & \text{otherwise.} \end{cases}$$

Each night contributes one wind and one footprint (§2.6), so a receptor accrues one detection encounter per night it is covered on, not one per pass. We count a night as one encounter rather than two because its two passes run minutes apart under the same wind and past the same buildings, so the misses $p_{\mathrm{enc}}$ represents are largely common to both and the passes are strongly correlated rather than independent (§3.1); the geometric benefit of driving both sides, which can reach pipe one side alone would miss, is already captured in the night's footprint. These encounters compound in closed form: a receptor covered on $m$ of the surveyed nights is detected with probability $1-(1-p_{\mathrm{enc}})^m$, with $m \le 3$ at six passes. Because this is the exact expectation over the per-encounter Bernoulli it adds no Monte-Carlo noise, and $p_{\mathrm{enc}} = 1$ recovers the pure-geometry union. Rather than making a particular assumption about $p_{\mathrm{enc}}$, we sweep values across $\{0.75, 0.86, 0.95, 1\}$ with $0.86$ serving as our "base" parameter for most of the analysis here, which is informed by the derivation in §2.5.

Finally, the base dispersion model we employ lets a plume pass straight through buildings. Because this is the standard way to generate coverage maps, and the basis of the single-pass coverage calculation we compare against (Tan and Rella, 2017), we report results from it by default: it accounts for the wind, but not for the obstructive interference of manmade infrastructure, in the coverage statistics. As an extension to this approach, we also apply a line-of-sight variant: a binary 2-D ray-march that blocks a receptor–sensor pair when an OpenStreetMap (OSM) building footprint sits between them. While this is a heuristic approach to account for obstructions, it is computationally tractable and a rough stand-in for the concentration deficit downwind of an obstruction. Figure 2 shows the effect on a single residential block: at one night's wind, roughly a third of the pipe the naive model counts as covered sits in a building's shadow and drops out. This model is coarse: it does not account for building heights nor reflection off walls (Micallef and Micallef, 2024), and it ignores channeling, where buildings funnel the near-surface flow along the streets and rotate the downwind plume axis by an additional 5–14° for winds oblique to

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

the street grid (Pirhalla et al., 2021). We make no claim that it is exactly right, but our conclusions do not depend on getting it exactly right. They rest on the difference between wind-blind and wind-aware surveying rather than on any absolute coverage level, and the more realistic maps only widen that difference: the harder detection becomes, the more sampling complementary wind direction matters. The obstruction-unaware (no-LOS) maps we use for the majority of the analysis are therefore the conservative choice for our claim. We quantify the effect of the more realistic maps in §3.3 and §3.4 (Figure 7), where the blind-versus-aware coverage gap is larger under the LOS-blocking plume model.

### 2.5 Pass-correlation test on public data

Two of our results draw on a public dataset rather than our own drive data: the pass-correlation analysis and the $p_{\mathrm{enc}}$ estimate. Luetschwager et al. (2021) provided openly accessible per-location drive-by detection records for 12 cities: for each leak location, the ordered sequence of drive-bys and whether each one detected the leak or not. Every sequence begins with a detection, since a location enters the dataset only by being detected once; that first entry is not an ordinary trial, so we drop it and analyze the drive-bys that follow, keeping locations with at least five drive-bys (and, for the order test, dropping the few whose remaining outcomes are all alike, which carry no ordering information). If the passes at a location were independent, the order in which detections and misses fall in that location's sequence would carry no information, and any reshuffling of the sequence would look as ordered as the real one. So we compare each location's real sequence against many random reorderings of its own detections and misses, holding the number of detections fixed. Because every location is only ever compared to reshuffles of itself, the test is insensitive to differences between locations in leak size, soil, city, or road speed; the only thing it responds to is the order of detections and misses. The drive-by sequences carry no timestamps, so we cannot recover how far apart in time successive passes fell; we therefore report the correlation in units of passes rather than hours.

The same records also give a plausible value for the residual detection probability $p_{\mathrm{enc}}$ of §2.4. Model a location, on a given run, as either feasibly detectable (a fraction $c$ of runs, each drive-by detecting with probability $p_{\mathrm{enc}}$) or not (detection $\approx 0$), with the state holding across its passes. The model has two unknowns, $c$ and $p_{\mathrm{enc}}$, and the records supply two numbers that fix them: the per-drive-by detection rate ($cp_{\mathrm{enc}} \approx 0.20$)[1] and the pass-to-pass agreement ($1 - 2cp_{\mathrm{enc}}(1 - p_{\mathrm{enc}}) \approx 0.95$; the §3.1 order test reports $90\%$ because it drops the always-agreeing all-miss locations kept here). Solving the two together gives $p_{\mathrm{enc}} \approx 0.86$ and $c \approx 0.22$, with $p_{\mathrm{enc}}$ staying within $0.85$–$0.87$ across reasonable choices of the location set and drive-by cutoff. We do not take $0.86$ as a precise value. The latent state is not shown to be wind (the authors' own caveat), the fit assumes one $(c, p_{\mathrm{enc}})$ for every location, and because the records carry no timestamps they mix same-night and different-night pass pairs, which can only lower the observed agreement and so makes $0.86$ a mild lower bound. This is why we sweep $p_{\mathrm{enc}}$ rather than fix it (§2.4); Appendix A works through the agreement and disagreement probabilities behind the estimate.

[1] This is the unconditional rate over all locations; on those ever re-detected it is about $0.34$, which under independence reaches $90\%$ in the five-to-eight drive-bys of the convention (Luetschwager et al., 2021).

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

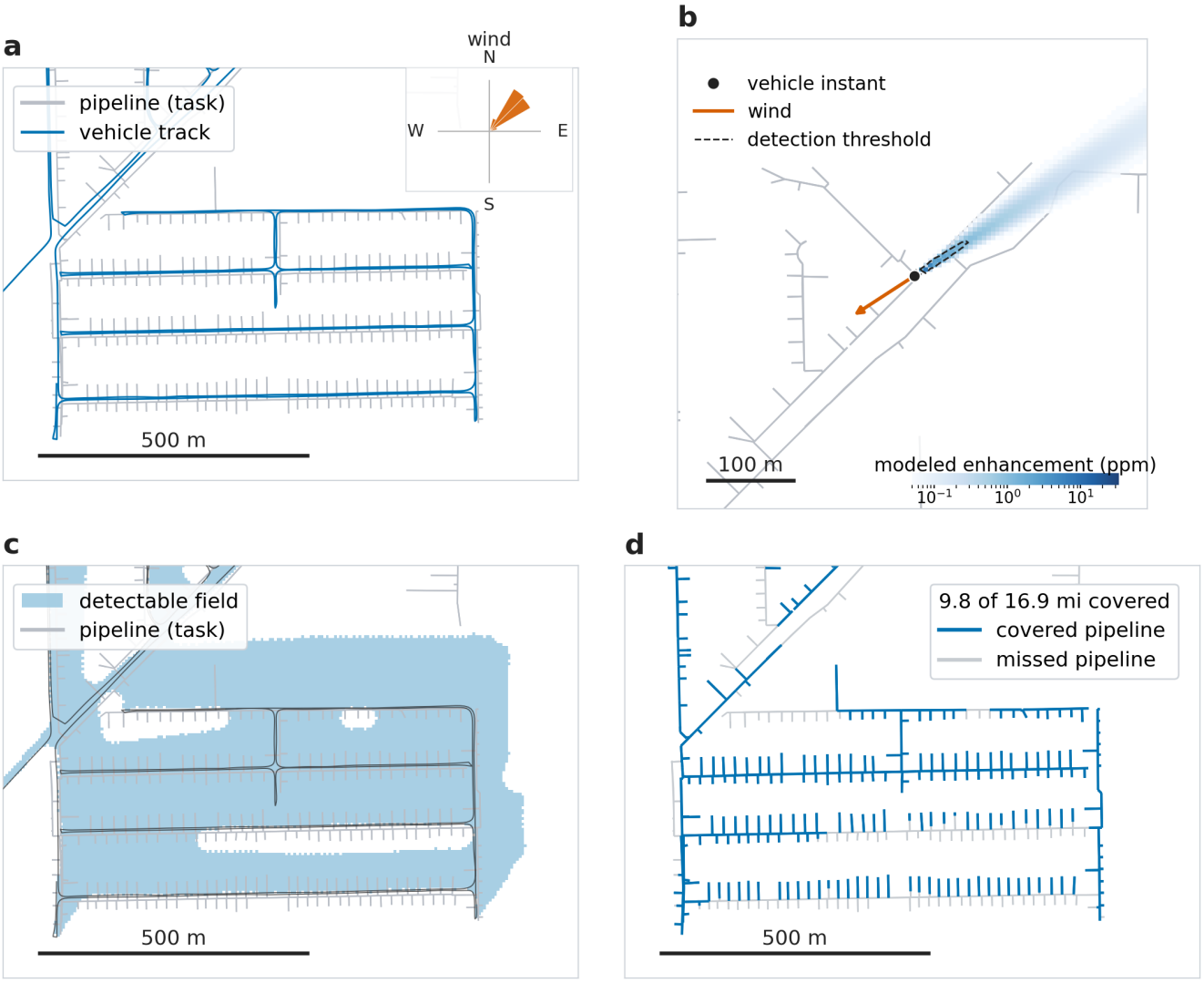



**Figure 1.** Computing the wind-conditioned coverable set for one survey drive. (a) Inputs: the pipeline the survey was tasked to cover, the vehicle track, and the night wind (rose, direction the wind blows from). (b) The inversion at a single vehicle instant. Shading is the methane enhancement a unit leak at each location would produce at the vehicle under the measured wind; the dashed line is the detection threshold. (c) The union over every instant, thresholded at the detection level at sensitivity $s = Q/\mathrm{MDL} = 360$: the drive's detectable field. (d) The detectable field rolled onto the pipeline, split into covered and missed, here 9.8 of 16.9 miles of pipeline at $s = 360$.

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

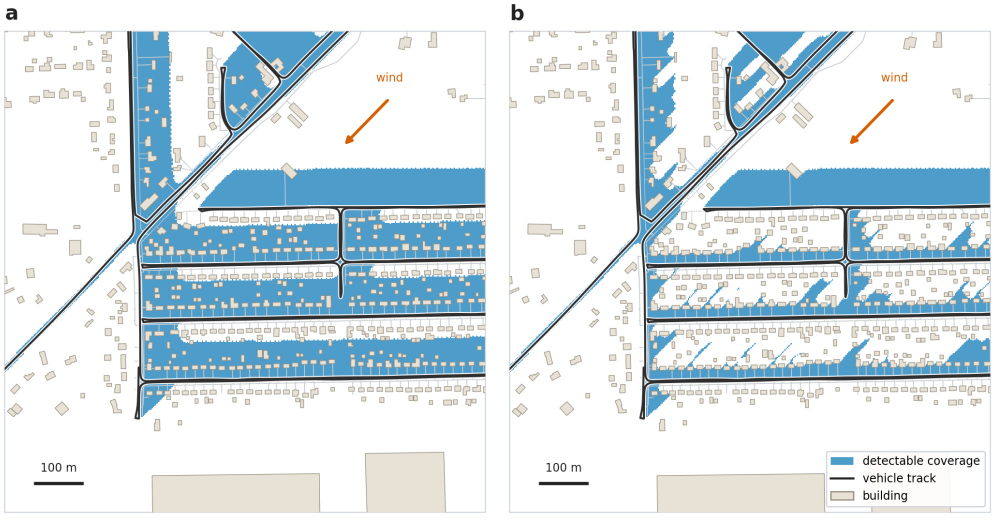



**Figure 2.** Effect of the line-of-sight correction on one residential block, at a single representative night wind (blowing from the northeast at 1.4 m s$^{-1}$). Blue is the thresholded coverable field, the region a leak could have been detected on; the black line is the vehicle track and tan footprints are OSM buildings. (a) Naive coverage, in which plumes pass freely through buildings, leaving a nearly continuous field. (b) With line of sight; a building between a potential source and the vehicle blocks the view, carving the field out on the leeward side of each building.

### 2.6 Monte-Carlo pass protocols

With the coverable set in hand for each drive and each candidate night, we compare two ways of accumulating passes. A night of surveying drives both sides of the street, which we count as two passes, so two, four, and six passes correspond to one, two, and three nights. The two passes of a night are the two sides of one short drive, of order $\sim$ 10s of minutes apart under essentially the same wind (the measured direction varies by a median of only about 13° over a drive, §3.2), so we give a night a single wind draw and one footprint rather than a separate draw per side; same-night re-drives separated by $\sim$hours, which do sample the within-night wobble, are treated separately in a sequential-nights check (§3.4). These pass counts are simulated, not observed: these are not coverage maps of any real multi-pass survey. Each drive contributes its route and network geometry once, and the repeat passes are generated by replaying that route under winds drawn for candidate nights from the ERA5 record at the drive's location. The coverage a protocol reaches is therefore a property of the scheduling rule applied to real routes and real regional winds, answering "had this route been driven on these nights, what would it have covered," rather than a measurement of how any particular survey in our data performed. The blind strategy mimics current practice: one trial draws its nights from a single seven-day window (the deadline-driven cadence operators describe), with the window itself placed at random in the multi-year record, so that repeating the trial samples the range of weeks a survey might land in. A trial's covered pipe is the union of the drawn nights' coverable sets, and its coverage is that union as a fraction of total pipe length, each segment counted in proportion to its length; we run 400 trials per drive. From those trials we summarize each drive by its own within-week coverage distribution, then report two spreads across the pool: the site-to-site spread of per-drive

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

medians (how coverage varies from one network to the next) and the typical network's own week-to-week spread (Figure 5); the two answer different questions and are kept separate throughout. The wind-aware strategy is deliberately an upper bound: it chooses the survey nights with hindsight. Seeing every candidate night's wind, a greedy rule adds, pass by pass, the night whose wind direction covers the most still-uncovered pipe, selecting on the winds that actually occurred rather than on the forecasts an operator would schedule from. That hindsight is exactly why the resulting curve is an upper bound rather than an achievable target: a real operator must commit to nights from an imperfect forecast, so the coverage actually realized sits below what selecting on the winds that did occur attains. We expect that gap to be small over the one-to-two-week horizon a survey plans within, since forecast error should degrade wind-aware coverage only modestly, and we leave a full account to future work. We run this Monte-Carlo simulation over a robustness grid of both plume models (naive and LOS), three sensitivities ($s \in \{90, 360, 1440\}$), and four residual detection probabilities ($p_{\text{enc}} \in \{0.75, 0.86, 0.95, 1\}$), so that no conclusion rests on a single choice of these parameters.

## 3 Results

### 3.1 Correlation of successive passes

Using the publicly available per-location drive-by records described in §2.5, we begin with the independence assumption. If each pass at a location were an independent try at catching the existing leak, then knowing that one pass succeeded would tell us nothing about whether the next one will. That is not what the data show. A location detected on one pass is detected again on the very next pass far more often than independence allows: successive passes agree about 90% of the time, where reshuffling each location's own passes at random gives only 63%. Measured against the spread of the reshuffled sequences, the observed agreement sits roughly two hundred standard deviations above what independence would produce (Figure 3). Independence is a natural simplification to make in the absence of useable wind data, but it is not a close approximation to what these records show.

The excess agreement appears in every one of the 12 cities on its own, and it survives every robustness check we ran: splitting locations by how many times they were driven, weighting each location equally, and testing each half of a location's sequence separately (per-city separations range from 23 to 109 standard deviations above chance). One further check points at the cause: the excess agreement between passes fades as the passes get further apart in the sequence and vanishes once they are about eight apart. That is the fingerprint of a condition that lasts for a handful of passes and then turns over, rather than a leak that is permanently easy or permanently hard to catch (agreement would never fade) or a set of independent passes (there would be no excess at any separation). This mirrors the temporal correlation Luetschwager et al. (2021) themselves report for plume magnitude, which fades over a few hours; they recommend spacing passes so as to sample a leak under varied conditions, an instinct our result supports and extends from plume size to detection itself.

We read this as picking up a thread the original study left open. The authors named the wind limitation themselves and anticipated that wind-aware surveys would be able to quantify spatial coverage; curious how far the independence assumption

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

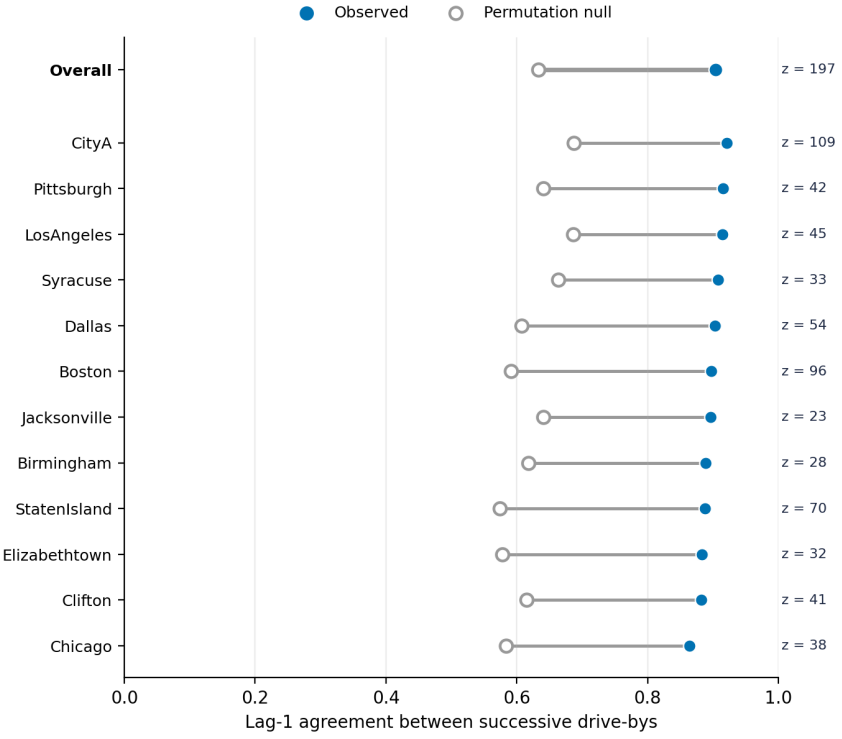



**Figure 3.** Lag-1 agreement between successive drive-bys vs. a within-location permutation null, overall and by city. Data of Luetschwager et al. (2021); one city is anonymized as CityA in that source.

sits from their own records, we measured it, and the way the agreement fades with separation points toward the mechanism they could not resolve without wind: a condition that holds for a few passes and then changes.

### 3.2 Two timescales of wind-direction variability

In order to understand what might cause the autocorrelated detection statistics of Luetschwager et al. (2021), we turn to our own drive data with on-board wind measurements. The lag structure of the previous section points to wind conditions that change over several passes, and the underlying driver, wind direction, varies on two very different timescales. Within a single night it is fairly stable, varying little over the course of a survey: across our drives the circular standard deviation of the measured direction over a pass has a median near 13°, tight enough that the second pass of a night points the plume the same way as the first. This is why, in these drives, driving at different hours of the same night rarely adds coverage to unobserved pipe: the direction usually has not changed enough to reach it. The change from one night to the next is gradual as well, so nights close together tend to resemble one another; the spread of directions widens only over a longer span, its circular standard deviation reaching ∼77° across a week and ∼116° across a year, by which point essentially every direction has occurred on some night even though a prevailing direction still dominates (computed across the full pool of drive sites). In our predominantly mid-

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

latitude data, then, the directions a survey can choose among broaden over days and weeks rather than within a single night or across a run of consecutive ones; how fast they broaden depends on the local wind climate and is not universal, a point we return to with the regional comparison in §4.

Figure 4 shows the consequence on one drive. A single night's wind produces a wind-conditioned coverage area spanning a median of about half the pipe (51%, naive, $s = 360$), and its shape is set by the direction the wind blew: the detectable pipe forms bands reaching upwind of the streets the vehicle drove, all pointing the same way. Re-driving under the same night's wind, whether an hour later or on a separate pass, re-covers those same bands and adds little. Three nights drawn blind from one week frequently repeat a direction and stack their coverage areas; three nights chosen for complementary directions tile the map instead. The correlation of §3.1 and the redundancy of blind repetition are both consequences of this effect.

### 3.3 Distribution of coverage at a fixed pass count

Because which pipe a pass covers depends on that night's wind, the coverage a fixed number of passes attains is not a single number but a distribution, and that distribution has two distinct sources. Coverage varies from one network to the next, since road and pipeline layout and local wind climate differ between networks; and it varies from week to week at a single network due to the distribution of wind speeds and directions. Figure 5 separates the two at six blind passes, under the naive, no-LOS physics that matches how single-pass coverage is normally reported, and at our default sensitivity $s = 360$. The two sources share the same center, a median of 81.0% of pipe, and turn out to be comparable in size: the middle 80% of networks span 66 to 89% at their typical week, while a single network's own coverage swings from about 65% in an unlucky week to 90% in a lucky one. Even at the median, and in this most favorable no-blockage case, on pipe the vehicle drove within 100 m of, the mobile survey reaches only about 81%, leaving nearly a fifth of the pipe for the hand-survey the protocol appends.

The low tail is concerning, and it is caused by both of the previously-mentioned sources (difficult-to-survey grids, and unlucky weeks): the hardest one in ten networks sits at 66% of pipe at its typical week, and any network caught in an unlucky one-in-ten week sits near 65% (naive, $s = 360$). Either way roughly a third of the pipe goes un-surveyed, and the tail deepens as the physics gets more realistic. Lowering the sensitivity to the operationally dominant $s = 90$ (the small-leak, high-threshold regime where real leaks concentrate, §2.4) pulls the whole distribution down to a median of 54.8% of pipe, with a low tail near 36% (naive, full pool). In that regime one network-week in ten leaves nearly two-thirds of the pipe un-surveyable. Adding line-of-sight occlusion, which removes pipe hidden behind buildings, lowers coverage modestly. Across the full pool (about 3,985 drives with building footprints, from local OpenStreetMap extracts), occlusion lowers the blind six-pass distribution by roughly four points at $s = 360$: the median falls from 81.0% to 77.2% and the tenth percentile from 66.1% to 62.5%, shifting the whole curve down rather than deepening only its tail. At the tighter $s = 90$ it barely moves (median 54.8% to 54.1%), since those footprints are already small enough that buildings rarely change what a plume reaches.

### 3.4 Choosing directions versus adding passes

If the redundancy of blind repetition results in sub-optimal coverage, then selecting nights with complementary wind directions should recover it at a fixed pass count. Figure 6 puts the two strategies side by side as a function of pass count. Blind driving

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

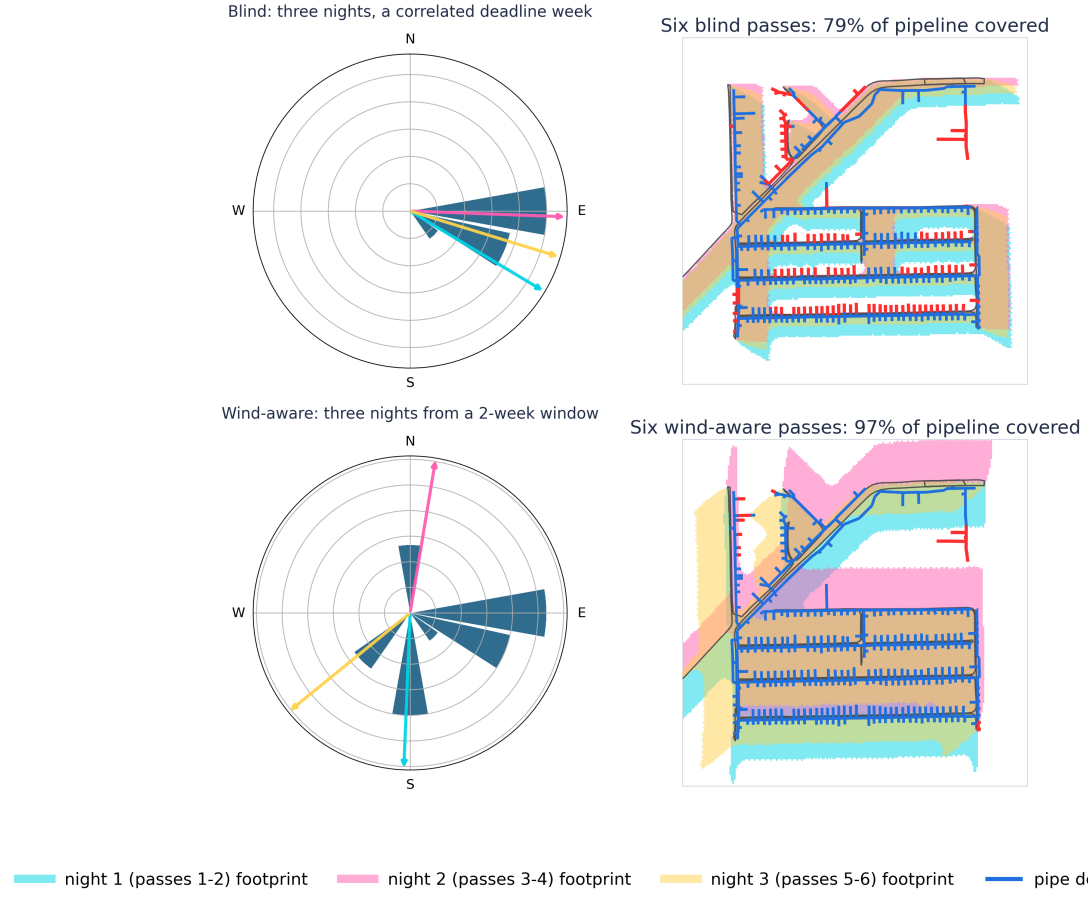



**Figure 4.** One survey drive (16.9 miles of pipeline), the same six-pass budget, two ways of choosing the nights. The top row is the blind case and the bottom row the wind-aware case; each row's left panel shows the three surveyed nights as arrows on that period's wind rose, and its right panel their detectable footprints, one color per night, rolled onto the pipeline (pipe detected in blue, missed in red). In the blind row the three nights cluster on nearly the same bearing and their footprints overlap; in the wind-aware row the nights are chosen from a two-week window for well-separated directions and the footprints spread across the route. Coverage is naive (no line of sight) at $s = 360$. The blind row shows one illustrative correlated week.

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

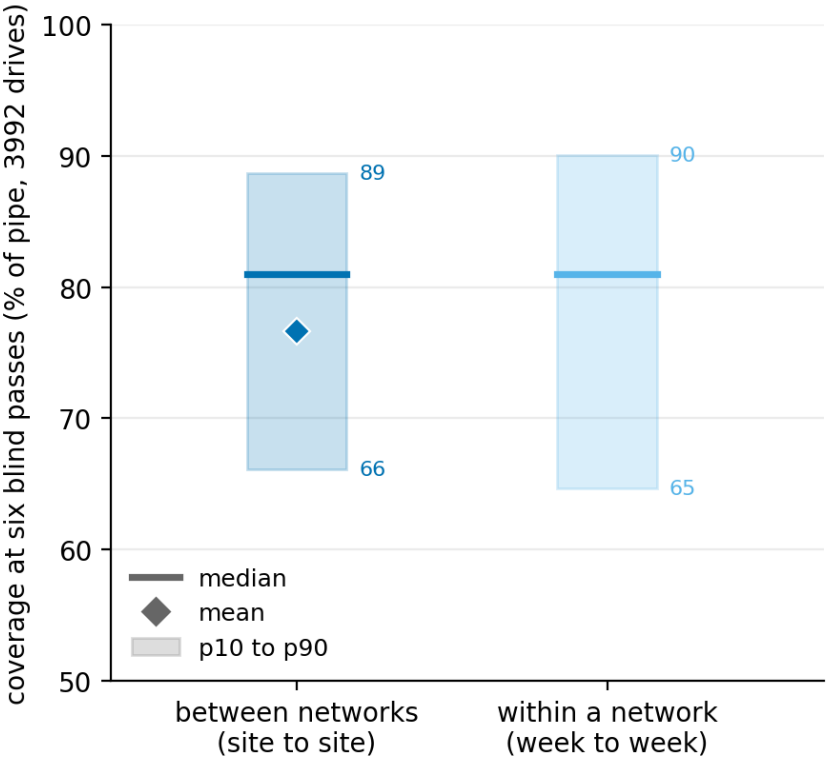



**Figure 5.** The two sources of spread in six-pass blind coverage of pipe (3,992 drives, naive, $s = 360$). *Between networks*: the across-drive spread of each network's typical-week (median) coverage. *Within a network*: the typical network's own week-to-week spread, the median across drives of each drive's within-week percentiles. Bars span [p10, p90]; the line is the median, shared by the two by construction; the diamond is the between-network mean.

(within-week) climbs 50.9%, 71.0%, 81.0% of pipe at two, four, and six passes (median, naive, $s = 360$, 3,992 drives). The wind-aware strategy, choosing nights within a two-week window, climbs faster and higher, to 67.5%, 87.6%, and 93.1%, so that four chosen passes (87.6%) clear six blind ones (81.0%) outright. This ordering is not a fragile artifact of the default settings. Evaluated at the wind-aware upper bound (the best the full record offers, which a two-week window nearly reaches; §3.5), aware-4 beats blind-6 across the physics grid, by a mean of $+13.2$ pp (naive, $s = 360$) and $+19.2$ pp (naive, $s = 90$). Line-of-sight occlusion, now computed across the full pool, preserves the advantage: comparing the same drives with and without occlusion, the aware-four-minus-blind-six gap is $+13.8$ pp with occlusion at $s = 360$, slightly larger than without, and $+19.2$ pp either way at $s = 90$, because occlusion removes coverage from the blind and wind-aware schedules alike.

The extra passes add little because they are not extra winds. Each night the vehicle drives both sides of the street under that night's wind, and that wind varies little through the night, so coverage is set by the number of distinct nights a survey samples, not by the pass count. The passes axis of Figure 6 is therefore a nights axis: two, four, and six passes are one, two, and three nights of surveying.

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

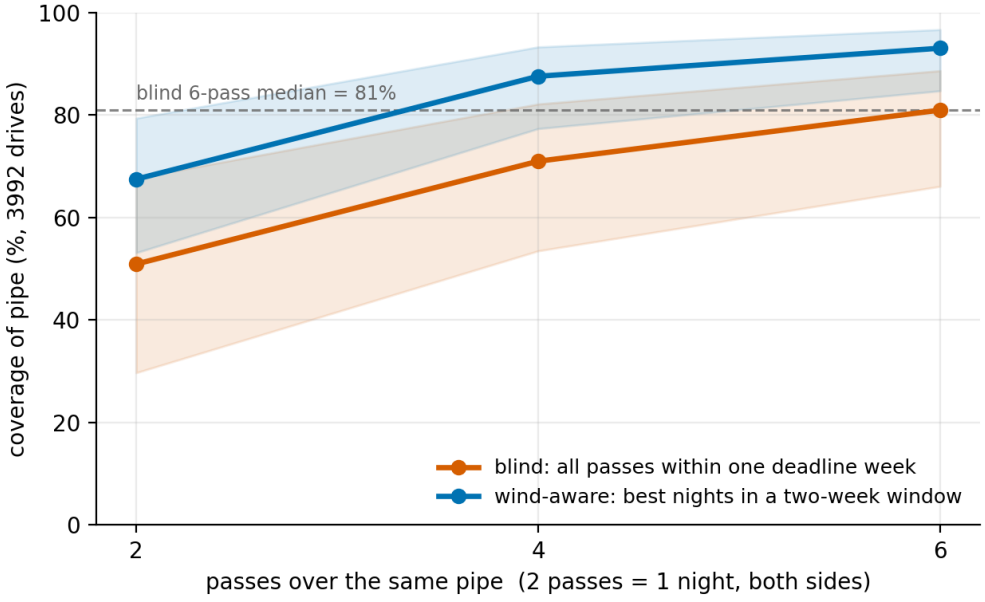



**Figure 6.** Coverage vs. number of passes, blind (within-week) vs. wind-aware (best nights chosen within a two-week window); median lines, [p10, p90] bands, dashed line the blind six-pass median, 3,992 drives.

The pass-count curves collapse each night to a single wind, so the two passes of a night share one footprint. A separate run relaxes this, drawing the two same-night re-drives at distinct hours within the work shift so they sample the real within-night wobble; it adds a little coverage, but far less than a fresh night would, so collapsing each night to one wind does not unfairly penalize the blind side.

The gap grows as the physics gets more demanding, which is the opposite of what a marginal effect would do. At six passes the aware-minus-blind mean gap runs from $+11.1$ pp in the easy corner (naive, $s = 1440$, where both strategies nearly saturate) to $+25.2$ pp in the hard one (naive, $s = 90$); adding line-of-sight leaves it at least as large in every column (full pool: the six-pass gap is $+25.2$ pp either way at $s = 90$, $+19.6$ pp with occlusion versus $+18.3$ pp without at $s = 360$, and $+13.0$ pp versus $+11.1$ pp at $s = 1440$), since occlusion lowers the blind and wind-aware coverage together (Figure 7). The reason is that smaller detection footprints make direction decisive: when the plume reaches only 25 m, pointing it the right way is most of what determines coverage, and added blind passes cannot reach pipe the night's wind never carried a plume across.

What matters operationally is the shape of the two distributions, not just their centers. The blind band here is the site-to-site one of §3.3: wide, with a long low edge at the hardest networks. The upper-bound band is narrow and high. Choosing by wind does less to lift the typical survey than to rescue that low edge, converting a distribution with a long low tail into one without; within a network, likewise, it removes the unlucky weeks.

Finally, the dominance survives the residual-detection sweep, which is where one might expect it to fail, since allowing per-encounter misses rewards driving the same pipe twice and blind repetition does exactly that. Applying the per-pass detection

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

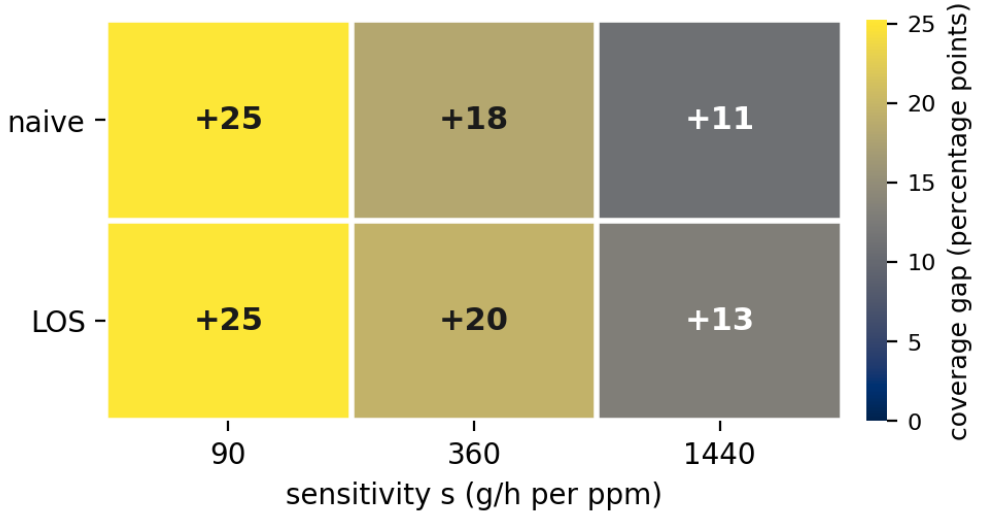



**Figure 7.** Blind-vs-chosen coverage gap across the physics grid ($\{\text{naive, LOS}\} \times s$), as the wind-aware minus blind mean coverage at six passes, in percentage points; cell values are means. Both rows are the full pool (naive $n = 3{,}992$, line-of-sight $n \approx 3{,}985$).

estimated from Luetschwager et al. (2021), $p_{\text{enc}} = 0.86$, to the full pool, the aware-4 versus blind-6 advantage (mean, naive) falls only from $+13.2$ pp to $+9.4$ pp at $s = 360$ and from $+19.2$ pp to $+15.4$ pp at $s = 90$, staying clearly positive in both. Sweeping $p_{\text{enc}}$ further at $s = 360$ traces a smooth, monotone decline that never reverses: $+13.2$, $+11.9$, $+9.4$, and $+6.3$ pp at $p_{\text{enc}} = 1$, 0.95, 0.86, and 0.75, so even at the most pessimistic detection rate the wind-aware schedule keeps a clear edge. The residual narrows the gap but never closes it, and it fails to close it precisely because it rewards redundancy. Since $1 - (1 - p_{\text{enc}})^2 > p_{\text{enc}}$, pipe that blind repetition re-covers on two nights of similar wind is more likely detected than pipe reached on a single night, whereas our distinct-direction schedule tends to cover each receptor on just one night. The reported aware curve is therefore a conservative floor for $p_{\text{enc}} < 1$; a planner that deliberately mixed a little redundancy with its direction diversity would only widen the advantage.

### 3.5 Sensitivity to the scheduling window

The wind-aware strategy is an upper bound, and the obvious objection is that it is unattainable: that choosing the right nights would require knowing the weather far in advance, or waiting arbitrarily long for a good direction to come around. Figure 8 answers by shrinking the window of directions the chooser is allowed to select from, and shows most of the benefit is available within a short window. Restricting the choice to the directions that occur within a single week, wind-aware six-pass coverage is already 90.8% of pipe (median); widening the window to a full year lifts it to 95.1%, essentially the 95.2% reached when nights may be drawn from the entire multi-year record (Figure 6), and most of that gain is captured within two weeks (93.1%). Four chosen passes clear six blind ones well inside the same window: wind-aware four-pass coverage is 85.3% at one week and 87.6% at two, already above the 81.0% six blind passes reach. Two-pass coverage moves from 65.5% to 70.7% over the

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

week-to-year range. The value of waiting is less in the median, which is largely settled within a week or two, than in the lower tail: the six-pass tenth percentile climbs from 77.5% at one week to 91.3% at a year, so a longer horizon mainly protects against the occasional week in which the available directions cluster.

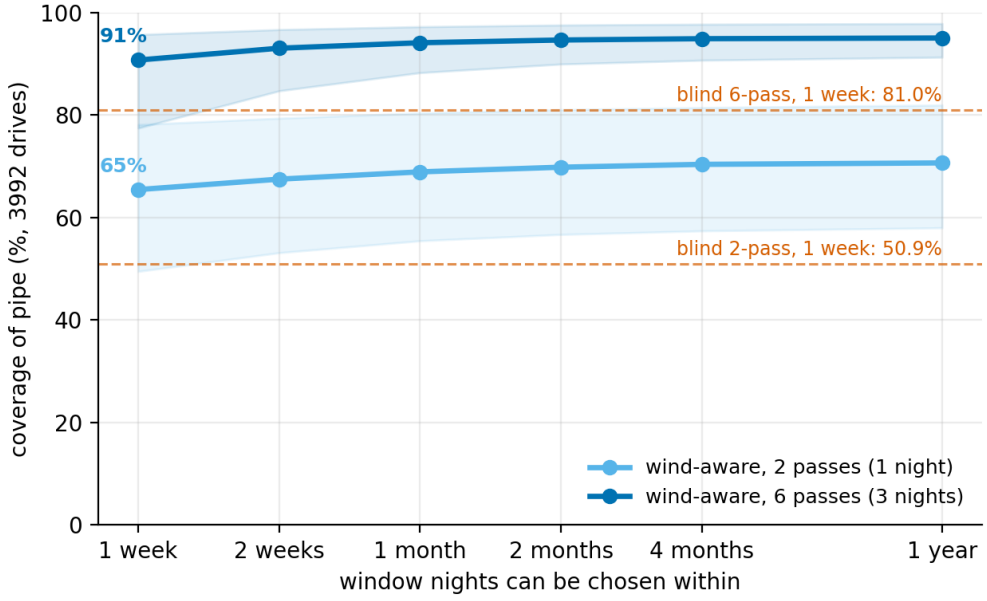



**Figure 8.** Wind-aware coverage of pipe against the length of the window within which nights may be chosen, at two fixed pass budgets (2 and 6 passes); lines are across-drive medians and shaded bands the [p10, p90]. Dashed lines mark the blind within-week median at the same two budgets.

The upper bound is therefore close to attainable. It is nearly reached by choosing among conditions that occur inside a one-to-two week window, a horizon short enough to schedule around with ordinary forecasts; closing the gap between the blind and wind-aware strategies does not require forecasting far ahead or waiting for months, which would be logistically at odds with how these systems are operated.

## 4 Discussion

The six-pass convention gets several things right. It identifies the variable that matters, since the diversity of wind directions is what makes repeated passes worth driving, and it validates against coverage, which is already what operators document and are judged on. Where it falls short is more subtle: the convention was calibrated to an average number of passes, but the coverage those passes produce varies from night to night, and simply counting passes does nothing to optimize coverage for the wind conditions on a given night. A week with clustered wind directions still yields the full six passes yet covers only a fraction of the pipe. Our results point to a different fix: more passes cannot recover coverage the night's wind never made possible,

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

whereas choosing passes by wind can, and that choice can be scheduled within a week and audited from the wind data the survey already records.

This also reconciles our finding with Luetschwager et al. (2021), which might seem to contradict it. The tension goes away because the two numbers measure different things: they estimate how many drive-bys capture >90% of *leaks*, while we characterize the distribution of *pipe* covered at a fixed pass count. Detection of a network leak factors into two terms, whether a length of pipe was in a position to be seen at all under the night's wind and, given a favorable pass, whether the instrument flagged the leak. Our coverage is the first of those terms and their per-pass detection statistic is the second, and the two multiply. Their own data even pin down the second: our two-state fit to those same records (§2.5) gives a per-pass detection probability of about 0.86, which sits inside the $p_{\mathrm{enc}}$ range we sweep. Multiplying a spatial coverage around 81% by a per-pass detection in that range reproduces the detection rates in our $p_{\mathrm{enc}}$ cells, so the two accounts agree.

Why, then, is our spatial coverage lower than the near-complete coverage a compliance survey is expected to deliver? In part because that near-complete figure is reached only after the protocol hand-surveys the coverage gaps the mobile passes leave (Picarro, Inc., 2019); our 81% is the mobile median before any such backfill, and it is the measured center of a wide distribution, reported as found. In part it is the denominator: our 100 m trace buffer is a proxy for the pipe a crew was tasked to survey, matched to work-order mileage to within about 5% but not identical to it, and any unassigned pipe it happens to include lowers the percentage. This affects the absolute level only, since the blind-versus-aware contrast behind our conclusions shares the same denominator in both arms and is insensitive to the choice.

The deeper issue is the independence premise itself. Treating each drive-by as an independent Bernoulli trial is a reasonable simplification in the absence of wind, and revisiting it in their posted data brings out the residual correlation it sets aside (§3.1); that correlation is, we argue, why an empirical calibration needed six passes in the first place: successive drive-bys under similar wind re-survey the same pipe, so each added pass buys less than an independent trial would.

Coverage is inherently local, though the data we have, concentrated in a few regions, cannot fully map how it varies across geographies. What a survey can cover, and how long it must wait to cover it, is set by how quickly the wind direction shifts from night to night in the local climate. There is a physical reason to expect the advantage of choosing to be largest where the range of night-wind directions is tightest: a regime that offers only a narrow range of directions is hardest on the blind protocol, because repetition keeps re-sampling nearly the same direction. Figure 9 illustrates the range across a handful of sites. At the few weather-mobile Midwest sites shown here the night-wind directions are widely spread (circular standard deviation about 113° over the record) and coverage climbs quickly toward saturation within a week, blind or chosen. These particular sites sit well above the pooled Midwest level: the pool mixes easier and harder networks, which is why its blind six-pass coverage centres on 81%, so these sites illustrate a wide-fan regime rather than represent the region. In a tropical Andes site (Bogota) the directions are tightly clustered (about 27°): blind six-pass coverage sits far lower, near 71%, and the wind-aware coverage climbs from about 82% at a one-week window to about 92% over a year, so the payoff to waiting is both larger and slower to arrive. The protocol and the physics are the same in both places; the answer differs because the local wind does. We show this to illustrate the mechanism. These regime groups are a small, hand-selected set of sites chosen to span distinct wind climates rather than a random regional sample, so they are not representative of any region's pooled coverage; with only two

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

to four drives per regime the comparison is underpowered, and a dedicated tail-risk test on the regime contrast does not reach significance. The absolute coverage level and the ordering among the wide-fan regimes depend as much on each site's road and pipe geometry as on its wind, so the Midwest sites sitting slightly below the Northeast and Gulf coast ones, despite a wider directional fan, should not be read as a wind-climate effect. The takeaway we do stand behind is the qualitative one, that coverage and the value of scheduling are case-by-case and climate-dependent, so a single headline number cannot generalize across regions. Assessing the relative impact of wind-aware scheduling on different climate regimes is deferred to future work.

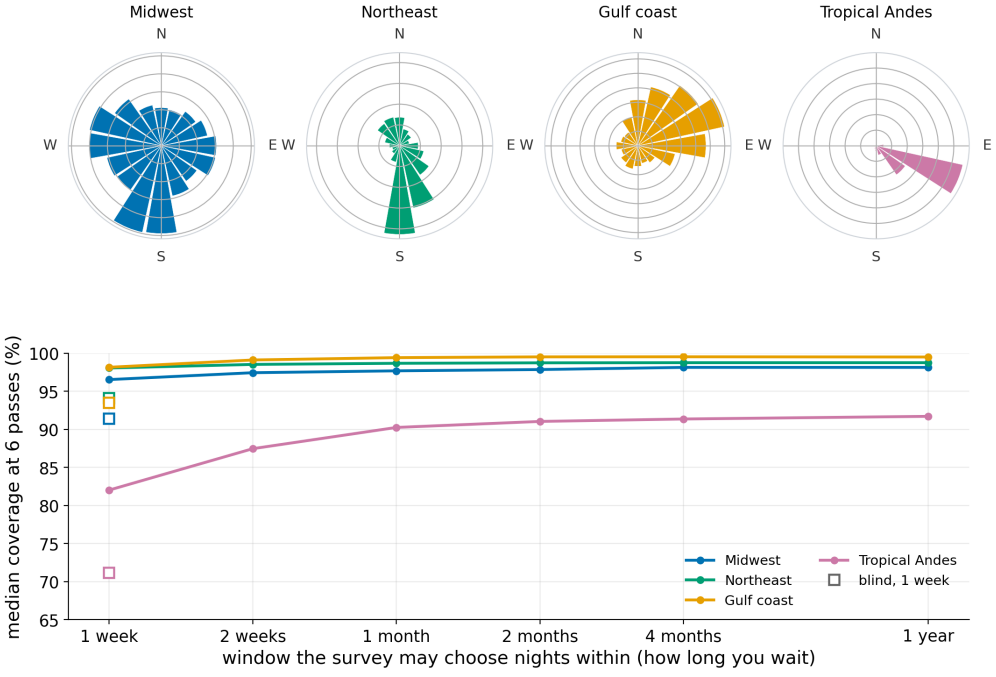



**Figure 9.** Night-wind climate and coverage across four climate regimes (illustrative; two to four drives per regime). Top: the distribution of ERA5 nightly-mean wind directions available at each regime's sites over the record. Bottom: wind-aware median coverage of pipe at six passes against the window within which nights may be chosen; open squares mark the blind one-week coverage.

The result bears directly on pending regulation. As of this writing the 2025 PHMSA leak-detection rule has been signed but has not appeared in the Federal Register, and its status there remains unsettled; it therefore has no effective date and is not yet codified. Its performance-based framing asks operators to tie detection equipment to the conditions under which it is valid and to a detection-performance standard, but it supplies no method for auditing either (Pipeline and Hazardous Materials Safety Administration, 2025). For a mobile survey the condition that matters is wind, and a wind-conditioned coverage computation is one such method: it turns "the conditions under which the survey could detect" from a claim an operator writes down into a number an auditor can recompute from the survey's own wind and track records, with no new instrumentation.

Our contribution draws on several literatures that have each addressed part of the problem. Stationary-sensor work optimizes placement over wind scenarios (Klise et al., 2020). Gao et al. (2022) bring wind into the routing loop by drawing a fixed-length

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

forecast downwind ray from each known point source to the nearest road; that approach plans ahead from a forecast, over discrete point sources, with simple geometry, where ours works backward from measured wind, over a continuous network, with a dispersion model. Mobile sensing interpreted through an atmospheric-transport model has likewise been used to detect and localize fugitive emissions across oil and gas production regions (Albertson et al., 2016), but over discrete sources rather than a continuous distribution network, and to find leaks rather than to measure and improve the coverage a survey achieves. The combination that is new here is the wind-conditioned coverable set of a continuous pipe network, computed under measured wind, used both to quantify what blind repetition leaves uncovered and to schedule survey nights that cover more. The same coverable set can also feed a close-enough arc-routing solver (Ávila et al., 2016) in place of a fixed footprint, letting a vehicle skip pipe that a given night's wind puts out of reach; in a 24-drive pilot this trims on the order of 12% of road-miles per drive, an operational payoff we note here but leave, with field validation, to future work.

It is worth noting several limitations and assumptions behind this analysis. Our coverage model is a near-field steady-state Gaussian plume, among the simplest and most computationally tractable dispersion models, and it carries the assumptions that come with that. We are comfortable with the choice for two reasons. The Gaussian plume is widely used and well studied, and, more importantly, our conclusions do not hinge on the absolute coverage any single model produces; what matters is the comparative difference between running the same model under different deployment strategies. We would not expect higher-fidelity dispersion modeling to change that difference, though confirming it is worth a look in future work. The line-of-sight variant is a coarse binary ray-march without building heights or wall reflections (Micallef and Micallef, 2024), and it ignores channeling. The wind-aware strategy is an upper bound, and while §3.5 shows most of it is reachable within a short window, we do not implement a forecast-driven planner. Our drives come predominantly from one mid-latitude US territory, with only a thin tail from other regions and climates, so the pooled coverage numbers describe that setting rather than a globally representative sample. Further limitations are a single-threshold detection rule (with a flat swept residual $p_{\mathrm{enc}}$ anchored to Luetschwager et al., 2021, whose latent state is not shown to be wind); a coverage denominator defined by a 100 m trace buffer, a proxy for the exact work order that matches its mileage to within about 5% but sets the absolute coverage level; and an ERA5 speed rescaling that affects footprint size but not direction.

## 5 Conclusion

Passes are not independent trials; their value lies in the conditions they sample, and because wind direction holds within a night, repeated passes under one night's wind re-survey the same pipe. This correlation is visible in the public data that has, in part, informed the six-pass convention. The consequence is that coverage at a fixed pass count is a distribution rather than a single number, set by how many distinct nights a survey samples: passes packed into one night largely repeat the same wind and re-survey the same fraction of pipe. Even at its median the distribution leaves a meaningful fraction of pipe uncovered, and below the median lies a lower tail, deepened by the smaller leaks that dominate real networks and pulled down further by line-of-sight blockage.

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Choosing passes by wind beats adding more blind ones. Four passes selected for complementary wind directions cover more pipe than six driven blind, in every combination of parameters and physics models we tested, and the margin widens as the leaks get smaller. Nearly all of that advantage is reachable within a one-to-two-week window of candidate nights.

*Code availability.* The code reproducing the pass-correlation re-analysis of the publicly available data from Luetschwager et al. (2021) (Sect. 2.5 and Appendix A) is available from the corresponding author on reasonable request. The remainder of the analysis depends on the proprietary survey data described below and is not publicly released.

*Data availability.* The survey data analyzed in this study are proprietary customer data held by Project Canary and cannot be made publicly available. De-georeferenced and de-identified data may be available from the corresponding author on reasonable request, subject to data-sharing agreements and the approval of the data-owning customers.

## Appendix A: Estimating the residual detection probability

This appendix works through the estimate of $p_{\mathrm{enc}}$ summarized in §2.5 and used in §2.4. We model each leak location, on a given run, as being in one of two states, with the state holding across that location's consecutive drive-bys: *feasible* (a fraction $c$ of the time), in which each drive-by detects independently with probability $p_{\mathrm{enc}}$, or *unfeasible* (the remaining $1-c$), in which detection probability is $\approx 0$.

The mixture has two unknowns, $c$ and $p_{\mathrm{enc}}$, and two quantities we can read off the records fix them. The first is the overall per-drive-by detection rate. A drive-by detects only if its location is feasible (probability $c$) and the encounter then succeeds (probability $p_{\mathrm{enc}}$), so the rate is $cp_{\mathrm{enc}}$, observed near $0.20$.

The second is the *agreement* between two successive passes at a location, the rate at which they return the same result (both detect or both miss). It is easiest to reason about through its complement, a disagreement, where one pass detects and the other misses. An unfeasible location never disagrees: both passes miss, so it always agrees. A feasible location disagrees only when its two independent drive-bys split, one detecting and one not, which happens with probability $2p_{\mathrm{enc}}(1-p_{\mathrm{enc}})$. A disagreement therefore requires both a feasible location *and* a split pair, with probability $c \cdot 2p_{\mathrm{enc}}(1-p_{\mathrm{enc}})$, so the agreement is

$$A = 1 - 2cp_{\mathrm{enc}}(1-p_{\mathrm{enc}}),$$

observed near $0.95$. Solving these two equations together, $cp_{\mathrm{enc}} \approx 0.20$ and $A \approx 0.95$, gives $p_{\mathrm{enc}} \approx 0.86$ and a feasible fraction $c \approx 0.22$. The estimate is stable: whether we include the always-miss locations or restrict to those with a mix of hits and misses, and whether we require two or five drive-bys, $p_{\mathrm{enc}}$ stays within $0.85$–$0.87$; only $c$ moves with the subset.

Two features caution against reading $0.86$ as a precise value. First, the fit assumes a single $(c, p_{\mathrm{enc}})$ for every location, and the latent state is not shown to be wind (the authors' own caveat). Second, the estimate is conservative in a predictable direction. Because the records carry no timestamps, the pass pairs we compare mix same-night and different-night pairs, and

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

across nights the wind, and hence a location's feasibility, can change. Since the model assumes feasibility holds steady across a location's passes, such a change appears as an extra disagreement, which lowers the observed agreement $A$ below its true same-night value. Holding the per-drive-by rate $cp_{\mathrm{enc}}$ fixed (it is a single-pass quantity, unaffected by how passes are paired), the agreement simplifies to $A = 1 - 2\,(cp_{\mathrm{enc}})(1 - p_{\mathrm{enc}})$, which increases with $p_{\mathrm{enc}}$; a lower observed $A$ therefore maps to a lower fitted $p_{\mathrm{enc}}$, so the true value can only sit above $0.86$. This is why we sweep $p_{\mathrm{enc}}$ in §2.4 rather than treat it as calibrated.

## Appendix B: Supplementary tables

This supplement collects the robustness checks the main text summarizes.

We first check the coverage denominator two ways. Comparing the 100 m trace buffer against the recorded planned-route (work-order) pipe on the 111 drives where a work order was logged, the absolute level differs as expected because the two count different pipe: on the work-order denominator the blind six-pass median is 94.1% against 81.0% on the buffer, since the buffer sweeps in unassigned pipe the crew was not tasked to survey. The comparative result is unchanged, with the wind-aware gain clearly positive on both denominators (aware-four-minus-blind-six $+9.4$ pp on the work-order set, $s = 360$, $p_{\mathrm{enc}} = 1$). Coverage also barely changes whether we count only the pipe labeled distribution-main or all pipe within the buffer: on the drives where both denominators were computed, the per-drive coverage gap correlates $r = 0.97$ across the two, so our results are not sensitive to that label.

Tables B1 and B2 give the full sensitivity sweep behind Figures 6–7: coverage of pipe versus pass count, blind and wind-aware, at $s \in \{90, 360, 1440\}$ under both plume models, reported as medians with [p10, p90] and as means. Table B3 reports the residual per-encounter detection sweep of §3.4.

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

**Table B1.** Coverage of pipe (%) versus pass count, blind versus wind-aware, across the sensitivity sweep (full pool, buffer denominator, $p_{\mathrm{enc}} = 1$). Median with [p10, p90] across drives. Line-of-sight occlusion lowers coverage more as the plume footprint grows, so it matters most at $s = 1440$; the blind-versus-aware ordering holds under both plume models at every $s$.

| Physics | $s$ | | 2 passes | 4 passes | 6 passes |
|---|---|---|---|---|---|
| Naive | 90 | *blind* | 26.8 [12.7, 34.8] | 44.0 [26.5, 53.5] | 54.8 [36.2, 65.4] |
| | | *aware* | 49.6 [38.7, 64.0] | 71.5 [58.4, 83.1] | 77.8 [64.3, 88.1] |
| Naive | 360 | *blind* | 50.9 [29.7, 67.8] | 71.0 [53.5, 82.2] | 81.0 [66.1, 88.7] |
| | | *aware* | 70.9 [58.3, 82.2] | 90.3 [84.2, 95.1] | 95.2 [91.7, 98.0] |
| Naive | 1440 | *blind* | 68.5 [44.8, 81.6] | 84.4 [69.6, 91.5] | 91.7 [82.8, 95.7] |
| | | *aware* | 81.3 [68.7, 89.6] | 96.1 [91.9, 98.5] | 98.8 [97.1, 99.9] |
| LOS | 90 | *blind* | 26.3 [12.5, 33.7] | 43.3 [26.2, 52.6] | 54.1 [35.8, 64.4] |
| | | *aware* | 48.8 [38.0, 63.4] | 70.8 [57.5, 82.5] | 77.3 [63.6, 88.0] |
| LOS | 360 | *blind* | 45.4 [28.5, 58.2] | 66.4 [50.5, 76.9] | 77.2 [62.5, 85.5] |
| | | *aware* | 66.9 [55.1, 79.4] | 87.4 [79.3, 93.9] | 93.2 [87.0, 97.4] |
| LOS | 1440 | *blind* | 60.0 [41.6, 73.2] | 79.3 [64.9, 87.6] | 87.7 [77.0, 93.5] |
| | | *aware* | 76.3 [63.6, 86.7] | 92.7 [85.7, 97.2] | 96.9 [92.7, 99.5] |

**Table B2.** Mean coverage of pipe (%) versus pass count, companion to Table B1 (full pool, buffer denominator, $p_{\mathrm{enc}} = 1$).

| Physics | $s$ | | 2 passes | 4 passes | 6 passes |
|---|---|---|---|---|---|
| Naive | 90 | *blind* | 26.8 | 42.0 | 51.7 |
| | | *aware* | 50.6 | 70.9 | 76.8 |
| Naive | 360 | *blind* | 49.6 | 67.5 | 76.6 |
| | | *aware* | 70.5 | 89.9 | 94.9 |
| Naive | 1440 | *blind* | 65.4 | 80.4 | 87.4 |
| | | *aware* | 80.1 | 95.5 | 98.5 |
| LOS | 90 | *blind* | 26.3 | 41.4 | 51.1 |
| | | *aware* | 50.0 | 70.3 | 76.3 |
| LOS | 360 | *blind* | 44.9 | 63.3 | 73.0 |
| | | *aware* | 67.2 | 86.8 | 92.6 |
| LOS | 1440 | *blind* | 58.3 | 75.3 | 83.3 |
| | | *aware* | 75.6 | 91.9 | 96.3 |

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

**Table B3.** Residual per-encounter detection sweep (naive, $s = 360$, full pool). Mean coverage of pipe (%) at four wind-aware and six blind passes, and their difference, as $p_{\mathrm{enc}}$ varies. The wind-aware edge narrows as redundancy is rewarded but never closes.

| $p_{\mathrm{enc}}$ | aware, 4 passes | blind, 6 passes | aware4 − blind6 |
|---|---|---|---|
| 1.0 | 89.9 | 76.6 | +13.2 |
| 0.95 | 87.1 | 75.2 | +11.9 |
| 0.86 | 81.6 | 72.2 | +9.4 |
| 0.75 | 74.1 | 67.8 | +6.3 |

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

*Author contributions.* DB performed the analysis and wrote the manuscript. NE, KO, and EC contributed to the methodology and to reviewing and editing the manuscript.

*Competing interests.* The authors are affiliated with Project Canary, which provides mobile methane leak-detection services.

*Acknowledgements.* Anthropic's Claude large language models were used to assist with editing and refining the text of this manuscript. The authors reviewed all content and take full responsibility for it. Road-network and building-footprint data from OpenStreetMap (© OpenStreetMap contributors, available under the Open Database License, ODbL) were used in the coverage analysis.

*Financial support.* This research was funded by Project Canary.

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

## References

Albertson, J. D., Harvey, T., Foderaro, G., Zhu, P., Zhou, X., Ferrari, S., Amin, M. S., Modrak, M., Brantley, H., and Thoma, E. D.: A Mobile Sensing Approach for Regional Surveillance of Fugitive Methane Emissions in Oil and Gas Production, Environmental Science & Technology, 50, 2487–2497, https://doi.org/10.1021/acs.est.5b05059, 2016.

Ávila, T., Corberán, Á., Plana, I., and Sanchis, J. M.: A new branch-and-cut algorithm for the generalized directed rural postman problem, Transportation Science, 50, 750–761, https://doi.org/10.1287/trsc.2015.0588, 2016.

Gao, M., Hugenholtz, C. H., and Barchyn, T. E.: Development and validation of a route planning methodology for vehicle-based remote measurements of methane and other emissions from oil and gas wells and facilities, Journal of the Air & Waste Management Association, 72, 1279–1289, https://doi.org/10.1080/10962247.2022.2113182, code: https://github.com/MozhouGao/DRIP, 2022.

Gundapuneni, V. R., Lo, J.-H., Zimmerle, D. J., and Smits, K. M.: Advanced Leak Detection Methods for Belowground Natural Gas Pipeline Leaks: Evaluation under Diverse Environmental and Operational Conditions, ACS Omega, 11, 7290–7302, https://doi.org/10.1021/acsomega.5c07766, 2026.

Klise, K. A., Nicholson, B. L., Laird, C. D., Ravikumar, A. P., and Brandt, A. R.: Sensor Placement Optimization Software Applied to Site-Scale Methane-Emissions Monitoring, Journal of Environmental Engineering, 146, 04020 054, https://doi.org/10.1061/(ASCE)EE.1943-7870.0001737, 2020.

Luetschwager, E., von Fischer, J. C., and Weller, Z. D.: Characterizing detection probabilities of advanced mobile leak surveys: Implications for sampling effort and leak size estimation in natural gas distribution systems, Elementa: Science of the Anthropocene, 9, 00 143, https://doi.org/10.1525/elementa.2020.00143, data: https://github.com/zdweller/ElementaDetectionProbabilities, 2021.

Maazallahi, H. et al.: Intercomparison of detection and quantification methods for methane emissions from the natural gas distribution network in Hamburg, Germany, Atmospheric Measurement Techniques, 16, 5051–5073, https://doi.org/10.5194/amt-16-5051-2023, 2023.

Micallef, A. and Micallef, C.: The Gaussian Plume Model Equation for Atmospheric Dispersion Corrected for Multiple Reflections at Parallel Boundaries, Sci, 6, 48, https://doi.org/10.3390/sci6030048, 2024.

Picarro, Inc.: Using the Picarro Asset Management Solution for Routine Regulatory Compliance Leak Survey of Distribution and Transmission Pipelines, Application Brief AB04-V1.1-191030, Picarro, Inc., https://www.picarro.com/sites/default/files/product_documents/Picarro%20Application%20Brief_Compliance%20Leak%20Survey.pdf, 2019.

Pipeline and Hazardous Materials Safety Administration: Pipeline Safety: Gas Pipeline Leak Detection and Repair (final rule as issued), https://pstrust.org/wp-content/uploads/2025/01/PHMSA-Final-Rule-Gas-Pipeline-Leak-Detection-and-Repair-As-submitted.pdf, docket PHMSA-2021-0039, RIN 2137-AF51; signed 2025-01-17; not published in the Federal Register and status there unsettled as of this writing; no effective date, 2025.

Pirhalla, M., Heist, D., Perry, S., Tang, W., and Brouwer, L.: Simulations of dispersion through an irregular urban building array, Atmospheric Environment, p. 118500, https://doi.org/10.1016/j.atmosenv.2021.118500, 2021.

Tan, S. M. and Rella, C. W.: Survey area indicators for gas leak detection, https://patents.google.com/patent/US9645039B1/en, u.S. Patent 9,645,039 B1 (Picarro, Inc.), granted 2017-05-09; family priority 2012-05-14, 2017.

Tan, S. M. and Rella, C. W.: Gas detection systems and methods using survey area indicators, https://patents.google.com/patent/US12235190B1, u.S. Patent 12,235,190 B1 (Picarro, Inc.), granted 2025-02-25; continuation of Ser. No. 16/450,970 in the chain from U.S. 9,557,240; claims priority to provisional 61/646,487 (2012), 2025.

https://doi.org/10.5194/egusphere-2026-6372
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



Tian, S., Riddick, S. N., Cho, Y., Bell, C. S., Zimmerle, D. J., and Smits, K. M.: Investigating detection probability of mobile survey solutions for natural gas pipeline leaks under different atmospheric conditions, Environmental Pollution, 312, 120 027, https://doi.org/10.1016/j.envpol.2022.120027, 2022.

Venkata Rao, G., Aldred, J. R., Zimmerle, D. J., and Smits, K. M.: Understanding detection success of belowground natural gas leaks in urban environments through controlled release experiments, Journal of Pipeline Science and Engineering, 6, 100 417, https://doi.org/10.1016/j.jpse.2025.100417, 2026.

von Fischer, J. C., Cooley, D., Chamberlain, S., Gaylord, A., Griebenow, C. J., Hamburg, S. P., Salo, J., Schumacher, R., Theobald, D., and Ham, J.: Rapid, Vehicle-Based Identification of Location and Magnitude of Urban Natural Gas Pipeline Leaks, Environmental Science & Technology, 51, 4091–4099, https://doi.org/10.1021/acs.est.6b06095, 2017.

Weller, Z. D., Roscioli, J. R., Daube, W. C., Lamb, B. K., Ferrara, T. W., Brewer, P. E., and von Fischer, J. C.: Vehicle-Based Methane Surveys for Finding Natural Gas Leaks and Estimating Their Size: Validation and Uncertainty, Environmental Science & Technology, 52, 11 922–11 930, https://doi.org/10.1021/acs.est.8b03135, 2018.