Leak detectability from pressure gauges in a water distribution network under correlated night-time noise, and the value of further gauges

Ricardo Marquez, Ph.D.
rmarquez123@gmail.com

September 2026

Abstract

A utility that already reads a few dozen pressure gauges must decide whether another gauge, or another kind of measurement, would let it see leaks it misses today. This estimation study answers that question for the 30 gauges of Areas A and B of L-Town, the network of the BattLeDIM benchmark, from the gauges' 2018 night pressures. The network's head-loss laws are linearized at a night operating point, which gives every gauge's pressure change per unit leak at every junction; the linear changes agree with EPANET in direction to a cosine of at least 0.955 and underestimate their size by a median 5 %. The night-to-night noise of a seven-night step statistic is modeled as a physical covariance, a night-demand scale plus regionally correlated demand deviations plus gauge noise, fitted by likelihood on 79 quiet windows. The fitted model captures where the noise lies (the dominant direction within 11 degrees) but not its size in the remaining directions: 29 of 30 whitened eigenvalues fall below and one above the band a correct model would produce. A validation fixed before it was run injected nonlinear 5 and 10 m^3/h leaks into the real 2018 pressures at 200 places and times: a detector told where and when the leak starts recovered 95.0 % and 98.5 % against predicted 90.6 % and 99.0 %, and the scan over all junctions raised 2 alarms in 11 quiet weeks against 0.21 expected. The validation failed. The noise model does not predict detection on real 2018 pressures; placements are ranked by relative information only, and no statement is made on whether a leak of a given size would be detected. Relative results stand: removing the least valuable gauges costs little until about 20 remain, Area B's single gauge is the least valuable, ten added gauges gain 0.172 bits per leak on top of 2.112, the greedy picks agree with an exhaustive search over pairs, and the first picks are stable under resampling. Under the model the median smallest detectable leak would fall from 6.1 to 5.4 m^3/h with ten more gauges, and across plausible noise covariances it lies between 4.2 and 6.1 m^3/h today.

1 Introduction

Leaks in a water distribution network are found from what the network already measures: the flows entering each supply area, the pressures at a few dozen gauges, and the night, when consumption is lowest and a leak is the largest share of the flow. A small leak that the area's flow balance cannot separate from night-to-night variation in consumption may still change the pressures nearby, and whether it does depends on where the gauges are. Adding a gauge costs money and a site visit; the question a utility can ask of its own data is which of its gauges already carry the information, what another gauge would add, and whether a different measurement, such as metered demand, would add more. Sensitivity-based placement for leak isolation has a long history in the water literature

(Pérez et al., 2011), and in statistics the choice of sensors that maximize the information about an uncertain field is a classical design problem with a greedy solution whose quality depends on how close the criterion is to submodular (Krause, Singh and Guestrin, 2008; Das and Kempe, 2011).

What decides the answer on a real network is the noise. The night pressures at neighboring gauges move together, because consumption rises and falls across a whole area at once, and they drift from week to week; a leak must be seen against that correlated background, not against each gauge's own scatter. This study uses the year of 2018 pressure readings from L-Town, the network of the Battle of the Leakage Detection and Isolation Methods (Vrachimis et al., 2022), to build a physical model of that background, and checks it against leaks injected into the same readings, with EPANET 2.2 (Rossman et al., 2020) as the hydraulic reference throughout. It is an estimation study: the network is linearized and the noise covariance is fitted by likelihood, and the design quantities are computed from both.

TERRA is a general-purpose engine for probabilistic inference on physical networks. It treats a network as a graph on which the physics is written locally: a conservation law at every node, a constitutive law on every edge, and the boundary and component laws that close them, attached to the same graph. Each of these is a relation the state of the network must satisfy, and the engine assembles them and solves them together. Domains enter through plugins that supply their laws, units and file formats; the solving and the bookkeeping are shared. This study uses its hydraulics plugin to linearize L-Town's head-loss laws around a night operating point, with the pressure-reducing valves in their active regime, and to place gauges greedily under a correlated noise covariance. Section 3 states the framework and each of these constructs precisely.

The study asks four questions. How well does the linear model reproduce the network's response to a leak? What noise does a leak have to be seen against, and does a physical model of it hold up against the data? Which existing gauges matter, and what would more gauges, or a known night-demand scale, add? And do the absolute answers — how small a leak can be detected — survive a test on the real pressures that was fixed before it was run? Section 2 describes the network and the data, Section 3 the methods, Section 4 the results, Section 5 their limits and Section 6 the answers.

2 System and data

The network. L-Town (Figure 1) has 782 junctions, 905 pipes, two reservoirs, one tank, one pump and three pressure-reducing valves (PRVs), a valve that throttles its flow to hold the pressure below it at a setting. The network divides into three supply areas. Area A (659 junctions) is fed from the reservoirs through PRV-1 and PRV-2; Area B (31 junctions) hangs behind PRV-3; Area C (92 junctions) is fed from the tank, which the pump fills from Area A. A head is the water's energy per unit weight, in meters; a pressure gauge reads the head less the ground elevation, so the two change together. The study covers Areas A and B: the 690 junctions there are the places a leak may occur, and the 657 of them that carry no gauge and are not a valve outlet are the candidate sites for a new gauge. Area C, whose pressures follow the tank level, is left out.

The data. The benchmark publishes a nominal EPANET model of the network and a year of 2018 readings at 5-minute resolution. This study uses the 29 gauges of Area A and the one of Area B, the three flow meters on the area's inflows, and the public list of the ten leaks repaired in 2018 with their repair times. Each night is summarized by the mean of the 60 readings from 01:00 to 05:55. The readings are simulated by the benchmark's organizers from a more detailed model with realistic consumption and leaks; they are the closest public stand-in for a utility's telemetry. The

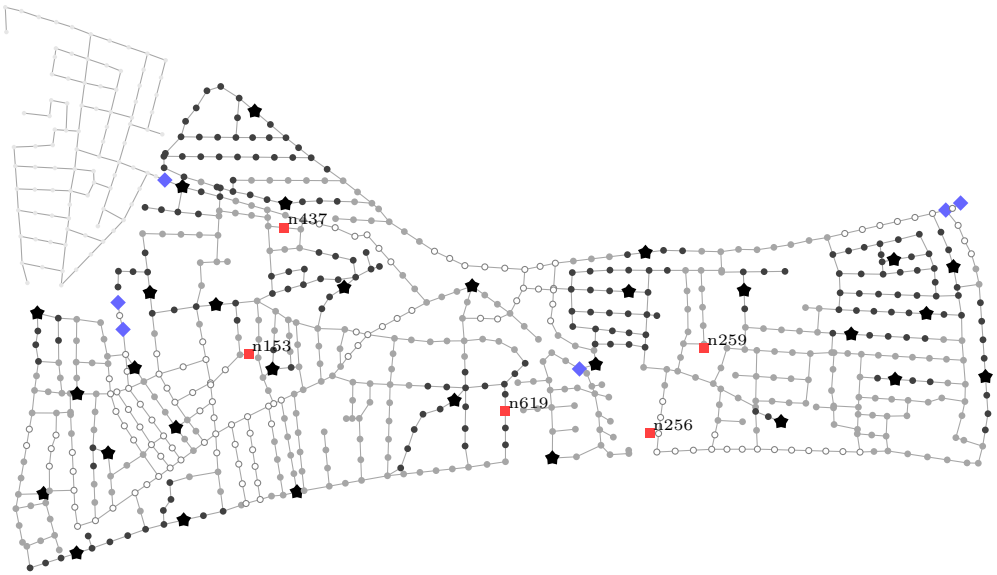


Figure 1: L-Town with its 30 Area A and B pressure gauges (stars), the three pressure-reducing-valve outlets, the two reservoirs and the tank (diamonds), and the first five added gauges of the greedy placement (red squares, labeled). Junctions of Areas A and B are shaded by the smallest leak the existing gauges would detect under the noise model of Section 3.4: dark at most 5 m^3/h, mid gray 5 to 8 m^3/h, open above 8 m^3/h; the pale dots are Area C, outside the study. Section 4 qualifies these values.

pipe roughness of Area A's three diameter classes is scaled by factors of 0.862, 1.066 and 0.923, fitted to 2018 pressures by an earlier calibration of the same network.

The operating point. At 03:00 of the demand pattern, EPANET puts all three PRVs in their active regime, passing 46.3, 51.8 and 4.6 m^3/h, and the pump running at 44.2 m^3/h. The nominal night demand of Areas A and B is 71.5 m^3/h, of which 26.0 residential and 45.5 commercial. A leak of 5 m^3/h is thus about 7 % of the night demand.

The questions. (1) How closely does the linearized network reproduce EPANET's response to a leak? (2) What is the covariance of the night pressure noise, and does a physical model of it fit? (3) Which gauges carry the information, and what do added gauges or a known demand scale add? (4) Do the absolute detection figures hold on the real 2018 pressures?

3 Methods

3.1 The linearized network

The unknowns are the head at every junction and the flow in every link. Conservation holds at every junction, and every open pipe obeys the Hazen–Williams law in the regularized form the hydraulic model uses,

$$\sum_{e \to i} Q_e - \sum_{i \to e} Q_e = d_i, \tag{1}$$

$$h_i - h_j = K_e \, Q_e \, (Q_e^2 + q_0^2)^{0.426} + K_e^{\mathrm{m}} Q_e |Q_e|, \tag{2}$$

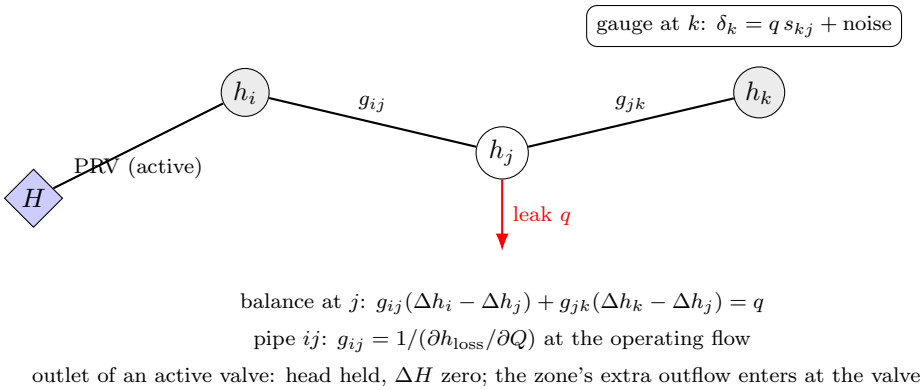


Figure 2: The linearized model around one junction. A leak is an extra outflow q at junction j. Each open pipe becomes a conductance g, the inverse slope of its head-loss law at the flow it carries at 03:00; the head change of every junction then follows from linear balances at the junctions, with the outlet of each active pressure-reducing valve held at its setting. A gauge at junction k sees the leak as q times the sensitivity s_{kj}, on top of the night-to-night noise of Section 3.4.

with Q_e the flow in m^3/s, d_i the junction's outflow, $K_e = 10.667 L_e/(C_e^{1.852} D_e^{4.871})$ from the pipe's length, roughness and diameter, K_e^{m} its minor-loss coefficient and q_0 a small flow that keeps the law smooth at zero. Reservoirs and the tank hold fixed heads; an active PRV holds its outlet head at its setting and passes whatever flow its downstream zone draws; the running pump follows its head curve. Figure 2 shows the linearization. At the 03:00 operating point every open link becomes a conductance, the inverse of its slope,

$$g_e = \Big(\frac{\partial (h_i - h_j)}{\partial Q_e} \Big)^{-1}, \tag{3}$$

taken from the same law the hydraulic model solves. The heads that are not held fixed then respond to an extra outflow q at junction j through the reduced conductance matrix G of the free junctions,

$$G\, \Delta h = -q\, e_j, \tag{4}$$

$$s_{kj} = \frac{\partial h_k}{\partial q_j}, \tag{5}$$

so that the sensitivity s_{kj}, the head change at gauge k per unit leak at j, is a column of $-G^{-1}$, found by one Cholesky factorization of G. Where a leak lies behind an active PRV, its zone's head is anchored at the valve outlet, and the extra flow the valve must pass reappears as an extra outflow at the valve inlet; chains of valves are solved together. The sensitivities are checked against EPANET solutions with a 5 m^3/h extra demand at 42 junctions.

3.2 The network model

The relations above are written as residuals on the network graph. A physical network is a graph, and the engine writes its physics on that graph. The state is a vector z of unknowns: one potential per node, such as a temperature or a head, one flow per edge, and whatever else the physics determines. Every relation the state must satisfy — a conservation law at a node, a constitutive law on an edge, a boundary condition, a component law — is written as a residual $r_k(z)$ that vanishes

when the relation holds. With as many independent relations as unknowns and no readings to reconcile, the state is the unique solution of

$$r(z) = 0, \tag{6}$$

found by Newton iteration on the sparse Jacobian J of the residuals, $J_{ki} = \partial r_k / \partial z_i$, from a stated start until the residual norm falls below a tolerance ϵ:

$$z \leftarrow z - J^{-1} r(z) \quad \text{until } \|r(z)\| < \epsilon. \tag{7}$$

A transient model adds a storage term to the conservation relations and advances the state in time with an implicit integrator, solving a system of the same kind at every step. Here the relations are solved once for the operating point and then linearized; the noise covariance of the next subsections is fitted to data by maximum likelihood, and the design quantities are computed from both. Everything that follows is a choice of unknowns and relations on this graph.

3.3 The step statistic

A leak that appears shifts the night pressures from then on. With y_n the vector of the gauges' night means on night n, the design statistic compares the latest seven nights with the seven before them,

$$\delta(n) = \tfrac{1}{7} \sum_{m=n-6}^{n} y_m - \tfrac{1}{7} \sum_{m=n-13}^{n-7} y_m, \tag{8}$$

which cancels anything constant over two weeks (a standing leak, a gauge offset, a roughness error) and the weekly cycle of consumption. A leak of size q starting on night $n-6$ adds $q\, s_j$ to $\delta(n)$. A step is *quiet* when no reported repair lies within 21 days of its first after-night; 79 of the year's steps are quiet. Leak starts are not public, so a quiet step can still contain one; the noise estimated from them is, if anything, too large.

3.4 The noise model

The noise of δ at any set of junctions is modeled physically. Night consumption at junction i changes between two weeks by $d_i \varepsilon_i$, with d_i its mean night demand and ε a relative deviation; the heads respond through the sensitivities, and each gauge adds its own error. The covariance is

$$\Sigma = \sum_c \theta_c \, (S d_c)(S d_c)^\top + \sum_a \theta_a \, u_a u_a^\top + s_r^2 \, S D R_\ell D S^\top + \sigma_s^2 I, \tag{9}$$

$$(R_\ell)_{ij} = \exp(-\mathrm{dist}_{ij}/\ell), \tag{10}$$

where S holds the sensitivities of the chosen junctions to every junction's outflow, d_c the night demand of consumption class c (residential and commercial; the industrial class has no demand in Areas A and B), u_a the head response to a unit change of PRV outlet a (a drift of the valve's setting), D the diagonal of the night demands, dist_{ij} the distance between junctions and σ_s the gauge noise. The parameters are fitted by maximizing the Gaussian likelihood of the quiet steps,

$$\hat{\theta} = \arg\min_\theta \ \log\det \Sigma(\theta) + \mathrm{tr}\left(\Sigma(\theta)^{-1} C\right), \tag{11}$$

with C the quiet steps' second moment and ℓ chosen from 100, 200, 400, 800 m. Because a step shares nights with every step within 13 nights, the 79 quiet steps hold only $N_{\mathrm{eff}} = 7$ independent

ones (the number of disjoint 14-night spans in the 5 quiet stretches of the year); N_{eff} scales the likelihood-ratio test of the regional term. The model is checked by whitening: if it were correct, the eigenvalues of $\Sigma^{-1/2} C \Sigma^{-1/2}$ would scatter as they do for data simulated from the model at the same nights, which gives a band per eigenvalue from 200 simulated years. The model is also refitted on non-overlapping steps, on resampled quiet stretches, on each half of the year, on same-weekday differences and with an 8-day guard, and asked to predict five held-out gauges from the other 25.

3.5 Detection, threshold and smallest detectable leak

For a leak at j the best linear test statistic of a step δ at the gauge set G is the matched statistic, which has unit variance without a leak and mean equal to the signal-to-noise ratio with one,

$$z_j = \frac{s_j^{\top} \Sigma^{-1} \delta}{\sqrt{s_j^{\top} \Sigma^{-1} s_j}}, \tag{12}$$

$$\mathrm{SNR}_j = q \sqrt{s_j^{\top} \Sigma^{-1} s_j}. \tag{13}$$

A leak must be found without knowing where it is, so the scan compares $\max_j z_j$ over the 690 junctions with a threshold z_{scan} set, by simulation of the noise, for a false alarm once a year at a weekly decision (probability $1/52$ per decision). A leak counts as detectable when it would cross that threshold nine times in ten, which gives the smallest detectable leak at each junction,

$$q_j^{\min} = \frac{z_{\mathrm{scan}} + 1.28}{\sqrt{s_j^{\top} \Sigma^{-1} s_j}}. \tag{14}$$

The medians and shares of $q^{\min}$ over the 690 junctions are the absolute results; their credibility rests on Section 3.7.

3.6 The value of a gauge

Placement uses a criterion that does not depend on a threshold: the information a gauge set carries about a leak amplitude of size $q = 5$ m^3/h at a junction drawn uniformly,

$$I(G) = \frac{1}{690} \sum_{j} \tfrac{1}{2} \log_2 \left(1 + q^2 \, s_{jG}^{\top} \Sigma_{GG}^{-1} s_{jG}\right), \tag{15}$$

in bits per leak. An existing gauge's value is the information lost when it is removed; the backward-elimination curve removes the least valuable gauge one at a time. Added gauges are chosen greedily, each maximizing the gain in I, which a rank-one update of Σ_{GG}^{-1} computes for all 657 candidates at once. Because the noise is correlated, I need not be submodular; the greedy pair is therefore compared with the best of all pairs, and the submodularity ratio, the sum of single gains over the joint gain of a set, is sampled on 500 random sets (Das and Kempe, 2011). The noise covariance at a candidate comes from the model (9). A second criterion counts, for a leak at each junction, how many other junctions it cannot be told from (their signatures within two whitened units), a measure of localization rather than detection.

3.7 Validation on the 2018 pressures

The absolute figures were tested by a rule fixed before it was run. Weekly decisions fall on nights $7k$ of 2018; 11 of them are quiet. 200 pairs of a junction (drawn uniformly from the 657 candidates)

and a quiet decision (drawn uniformly) were drawn with a fixed seed. For each pair and each size, 5 and 10 m^3/h, the nonlinear EPANET head change of a constant extra demand at that junction, at the 03:00 operating point, was added to the real night pressures from the first night of the decision's after-window on, so the decision sees a full step. Detector (i) knows the junction and tests $z_j \geq z_1 = 2.070$, the one-sided threshold for the same false-alarm rate; its predicted probability is $\Phi(\mathrm{SNR}_j - z_1)$ with the linear signature. Detector (ii) scans all junctions against $z_{\mathrm{scan}} = 3.476$ and counts only when the maximizing junction is within 300 m of the leak along the pipes; its prediction is $\Phi(\mathrm{SNR}_j - z_{\mathrm{scan}})$. The mean prediction over the pairs is compared with the recovered fraction and its 90 % Wilson interval (Wilson, 1927). Detector (ii) was also run at every quiet decision without an injection. The validation passes when detector (i)'s prediction lies inside its interval at both sizes and the false alarms do not exceed the 95 % Poisson bound of the design rate; otherwise the report states that the noise model does not predict detection and gives absolute figures only as a range across covariances.

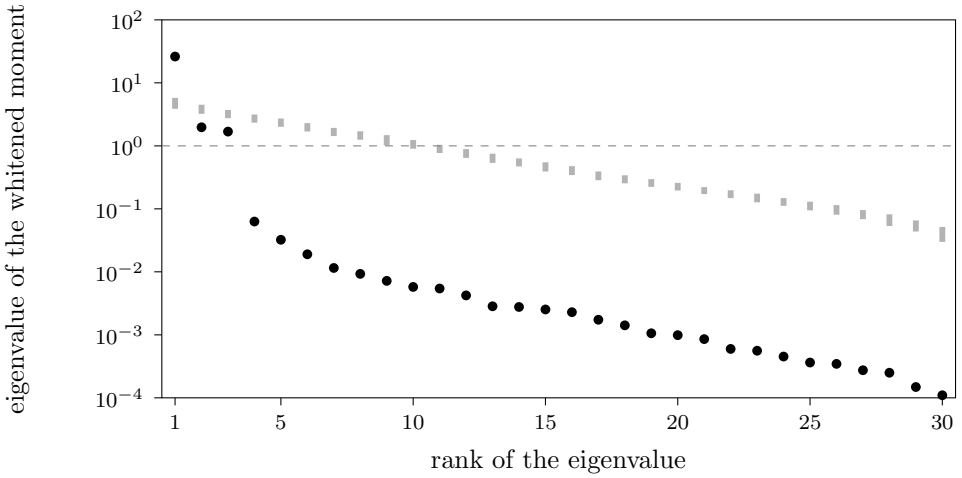
3.8 The 2018 repairs

The ten repairs of 2018 are the only known changes in the data. For each of the nine in Areas A and B, the step across the repair day (seven nights after less seven before) is compared with the step the model predicts for removing a leak of the size the flow meters show; the matched size $\hat{q} = -s^\top \Sigma^{-1} \delta / s^\top \Sigma^{-1} s$ is the size that explains the observed step best, with s the mean of the repaired pipe's two end junctions' signatures. The area's inflow over the same windows bounds any leak's size; EPANET at the measured size gives the nonlinear response; and the sensitivities are recomputed about a state that still contains the leaks repaired later in the year, which they did at the time. The two repairs four days apart are fitted jointly.

4 Results

Linearization. Of the 42 check junctions, 40 are seen by Area A's gauges (the others are an Area C junction and the outlet of PRV-1, next to a reservoir). With the gauges' common level removed, the linear and EPANET head changes at 5 m^3/h agree in direction to a cosine of at least 0.955 (median 0.9991), and differ by a median of 6.1 % of EPANET's change (90th percentile 14.1 %). The linear change is always the smaller, 0.82 to 0.98 of EPANET's: a finite leak raises the flows and so the slopes of the head-loss law. The difference is a finite-size effect, not an error of the slopes: at junction n106 it falls from 8.4 % at 5 m^3/h to 0.6 % at 0.25 m^3/h.

The noise. The quiet steps' second moment has one dominant direction of 200 mm standard deviation, then 23, 15, 5.1 and 2.2 mm, and a floor below 1 mm. The dominant direction is the response to a change of the night-demand scale (cosine 0.979): the whole area's consumption moving together. The fitted model puts the relative night-demand scale at a standard deviation of 19.6 % between two weeks, regionally correlated deviations of 30 % over $\ell = 200$ m, and gauge noise of 3.4 mm; the PRV outlets and the commercial class take no variance. The regional term is needed (likelihood ratio 172 on two degrees of freedom at $N_{\mathrm{eff}} = 7$). Figure 3 shows the whitening check: the model's dominant direction is 11 degrees from the data's, and that small angle, on a mode of 200 mm, leaves more variance than the model allows in one direction (whitened eigenvalue 26.2 against a band up to 5.8); the fit compensates with a gauge noise far above the data's floor, so 29 of the 30 eigenvalues fall below their band. The fit is stable under resampling of the quiet stretches (demand-scale variance 0.019 to 0.064, gauge noise 2.5 to 4.5 mm, 10th to 90th percentile) and



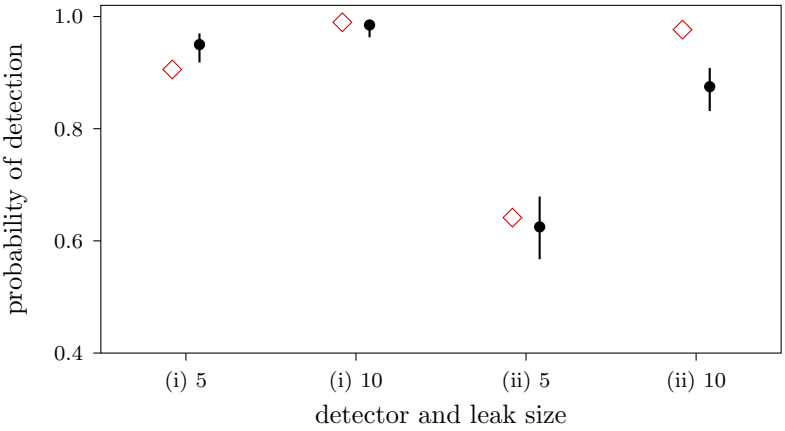
dots: the data whitened by the fitted model; gray bars: 5–95 % of the same quantity for data drawn from the model at the same nights

Figure 3: Whitening check of the noise model. The quiet-window second moment of the step statistic, whitened by the fitted covariance, should have eigenvalues inside the gray bars if the model were correct; the bars come from 200 simulated years with the same windows. 1 eigenvalue lies above its bar and 29 below theirs: the model places too little variance in one direction and too much in almost all others.

between the two halves of the year, but on same-weekday differences and with an 8-day guard the demand-scale term vanishes into the regional one at the shortest length. Predicting five held-out gauges from the other 25, the ratio of the observed to the predicted conditional variance runs from 0.05 to 79 (10th to 90th percentile over 12 splits): the model is right about new gauges on average and wrong for most of them individually.

Validation. Figure 4 compares prediction and recovery. At 5 m^3/h detector (i) recovered 95.0 % of the injections (interval 91.8 % to 97.0 %) against a prediction of 90.6 %, outside the interval; at 10 m^3/h it recovered 98.5 % against 99.0 %, inside. The scan, detector (ii), recovered 62.5 % and 87.5 % against predictions of 64.1 % and 97.7 %; 29 and 21 injections crossed the threshold but pointed more than 300 m away, which the prediction does not count as a miss. Without injections, the scan raised 2 alarms in the 11 quiet weeks (all at junction n54, the pump's suction), against 0.21 expected; the Poisson bound allows 1. With so few decisions this check has little power, and it still failed. The validation failed. The noise model does not predict detection on real 2018 pressures; placements are ranked by relative information only, and no statement is made on whether a leak of a given size would be detected. The absolute figures that follow are statements of the model, and Table 3 gives the range they span.

The 2018 repairs. Table 1 lists the nine repairs. In every case the repaired pipe ranks high among the 690 junctions by the matched statistic (ranks 4, 11, 6, 26, 1, 18, 2, 1, 16), and with the night-demand directions removed the observed and predicted steps still agree in direction (cosine 0.65 to 0.99). Their sizes do not all agree, and the table shows why. For p369 the flow meters' drop (5.9 m^3/h, three nights before against one after) and the inflow over the statistic's own windows (20.7) disagree, and the pressure size (20.2) sides with the latter. For p673 in Area B, p866 and p158 the pressure size exceeds the inflow step, which bounds any leak: 84.5 against 21.3, 36.2 against 20.1, 25.1 against 9.7. Two model effects enlarge them. EPANET's response at the measured size



filled circles and bars: fraction recovered from 200 injections into the 2018 pressures, with its 90 % Wilson interval;
open diamonds: the fraction the noise model predicts

Figure 4: Predicted against recovered detection for the two detectors of Section 3.7 at leaks of 5 and 10 m^3/h. Detector (i) knows where and when the leak starts; detector (ii) scans all junctions and must also point within 300 m of the leak.

is 1.03 to 1.36 times the linear one, and taking the sensitivities about a state that still holds the year's other leaks, as the network did then, lowers every matched size (p866 to 11.8, p673 to 33.5). Fitted jointly, the two repairs four days apart give 41.3 and 50.9 m^3/h against an inflow step of 47.8 over their combined span. The direction of a leak's signature is reliable; its size at a given leak depends on the state of the network around it.

The existing gauges. Table 2 ranks the gauges by the information lost when one is removed, and Figure 5 removes them one at a time. Under the model, the median smallest detectable leak with all 30 gauges is 6.07 m^3/h; removing the ten least valuable raises it only to 6.74, and the next ten to 8.27; below ten gauges it rises steeply, to 31.6 with one. Area B's gauge n215 is the first to go: behind an active PRV it sees only its own 31 junctions, while every Area B leak reaches Area A's gauges as the same extra PRV-3 flow. The gauges that matter most (n644, n613, n458) sit far from the valve outlets, where heads are least anchored.

The demand mode and a demand measurement. Table 3 repeats the existing gauges' figure under five covariances. The median smallest detectable leak ranges from 4.15 m^3/h, when each gauge's own scatter is taken as its noise, to 6.07 m^3/h under the fitted model; ignoring the correlation between gauges is the most favorable choice. Knowing the night-demand scale, as metered consumption could supply, lowers the model's median from 6.07 to 5.96 m^3/h, less than one added gauge does (5.94): in the fitted model most of the demand variance is regional, not area-wide, and a single scale does not remove it. Without the regional term the answer reverses, from 4.40 to 2.85 m^3/h: the value of a demand measurement is decided by how the demand noise is structured, which the 2018 data do not pin down (Section 5).

Added gauges. Table 4 lists the greedy additions. Each adds between 13.7 and 24.8 millibits per leak, less than each of the 10 most valuable existing gauges carries, and the median smallest detectable leak falls from 6.07 to 5.56 m^3/h with five and 5.36 with ten. The picks are not sharp: the runner-up is within 1.8 % of the best at every step, and 3 to 28 candidates are within 10 %, so

Table 1: The nine 2018 repairs in Areas A and B. *drop*: the night inflow before less after the repair from the flow meters (3 nights before, 1 after); *inflow step*: the same meters over the statistic's own windows (7 and 7 nights); *matched*: the leak size that best explains the pressure step at the gauges under the noise model; *re-linearized*: the same with the sensitivities taken about a state that still holds every leak repaired later in the year; *rank*: where the repaired pipe's junction ranks among the 690 by the matched statistic; *cosine*: agreement in direction of the observed and predicted steps once the night-demand directions are removed; *EPANET/linear*: the size of EPANET's head change at the drop against the linear one. Sizes in m^3/h.

pipe	area	repaired	drop	inflow step	matched	re-linearized	rank	cosine	EPANET/linear
p232	A	2018-02-10	32.0	22.2	24.2	8.5	4	0.95	1.18
p673	B	2018-03-23	11.4	21.3	84.5	33.5	11	0.95	1.15
p461	A	2018-04-02	38.2	29.5	34.0	15.6	6	0.98	1.36
p628	A	2018-05-29	36.9	37.9	39.6	20.1	26	0.91	1.22
p538	A	2018-06-02	27.0	32.4	41.4	20.9	1	0.65	1.19
p866	A	2018-06-12	18.3	20.1	36.2	11.8	18	0.99	1.35
p183	A	2018-09-01	23.9	13.2	20.7	7.1	2	0.99	1.27
p158	A	2018-10-23	26.2	9.7	25.1	14.8	1	0.94	1.21
p369	A	2018-11-08	5.9	20.7	20.2	16.1	16	0.99	1.03

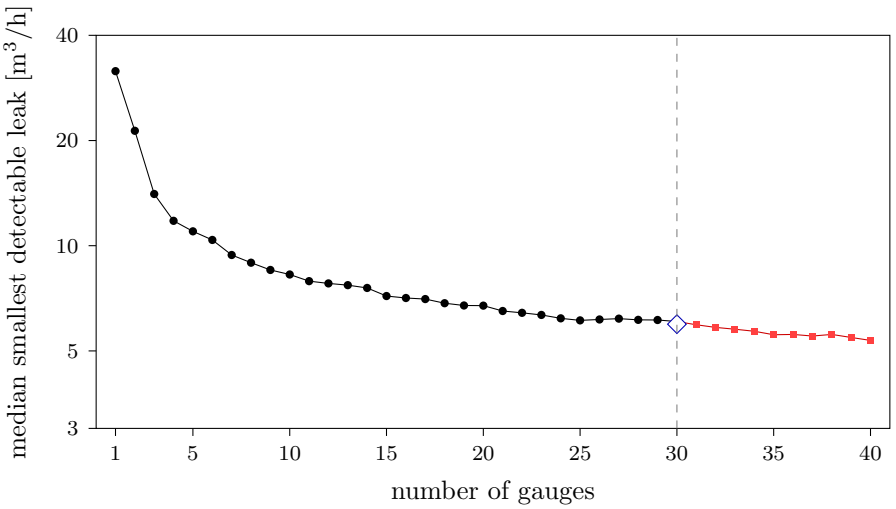
each pick marks a region rather than a junction. The best pair over all 657 candidates is the greedy pair (n259, n437); the submodularity ratio is at least 0.961 over 500 sampled sets and below one in 29 % of them, so greedy is close to optimal here without a guarantee. Refitting the noise model on 20 resampled years picks n259, n619, n153 among the first five every time; the half-year fits and a free level per zone give the same five, the same-weekday fit 3, the 8-day-guard fit 2, the half-data covariance 3, and the model without its regional term only 1. Placed for localization instead, five gauges (n600, n260, n190, n324, n621) cut the number of junctions a 5 m^3/h leak cannot be told from from 21.8 to 16.0, against 17.8 for the five detection picks; the two criteria choose different places.

Table 2: Information lost when one existing gauge is removed, largest five and smallest five of 30 (in thousandths of a bit per leak, criterion (15)). For comparison, the best single added gauge gains 24.8 millibits.

gauge	information lost [millibits]
n644	71.5
n613	56.7
n458	55.8
n296	42.6
n229	35.2
⋮	⋮
n429	14.1
n519	12.8
n495	10.0
n105	8.0
n215	5.9

Table 3: The smallest detectable leak with the 30 existing gauges under five noise covariances (m^3/h over the 690 junctions); *threshold* is the scan threshold z_{scan} of (14). The spread between rows is the range the absolute results carry.

noise covariance	median	10–90 %	share at 5 m^3/h	threshold
physical model (factors and regional term)	6.07	3.19–8.73	32 %	3.46
factors only, no regional term	4.40	2.95–7.21	67 %	3.34
half data, half model at the gauges	4.83	2.43–7.45	53 %	3.50
each gauge's own variance, no correlation	4.15	3.41–4.75	93 %	2.76
physical model, night-demand scale known	5.96	3.14–8.59	34 %	3.46



circles: existing gauges removed one at a time, least valuable first; squares: gauges added by the greedy placement; open diamond: the 30 existing gauges with the night-demand scale known

Figure 5: The median, over the 690 junctions of Areas A and B, of the smallest leak the gauges would detect, as gauges are removed from the existing 30 (left of the dashed line) and added to them (right). The curve is a model statement under the noise model of Section 3.4 (see Section 4); its steps are not all monotone because each gauge set has its own detection threshold.

Table 4: Greedy placement of up to ten gauges added to the existing 30. I: the information criterion (15); *median leak*: the median smallest detectable leak in m^3/h; *share*: junctions where 5 m^3/h is detectable (model statements, Section 4); *gap*: how much less the runner-up would have gained; *near*: candidates within 10 % of the pick's gain; *picked*: how often the node is among the first five picks when the noise model is refitted on resampled quiet segments.

k	added	I [bits]	median leak	share	gap [%]	near	picked
0	(existing)	2.112	6.07	32 %			
1	n437	2.137	5.94	34 %	0.61	12	17/20
2	n259	2.157	5.84	36 %	0.02	25	20/20
3	n619	2.178	5.77	38 %	1.14	17	20/20
4	n153	2.197	5.69	39 %	1.76	3	20/20
5	n256	2.212	5.56	41 %	0.49	18	13/20
6	n555	2.227	5.57	40 %	0.66	25	0/20
7	n655	2.242	5.52	41 %	0.63	21	4/20
8	n71	2.256	5.57	42 %	0.26	23	0/20
9	n83	2.270	5.46	43 %	0.30	26	0/20
10	n454	2.284	5.36	45 %	0.34	28	0/20

5 Discussion

The noise model is the weak part of the study, and the validation says so. A covariance built from the network's physics finds where the night noise lives — almost all of it in the response to consumption rising and falling together — but a few degrees of error on a mode of two tenths of a meter is enough to misplace more noise than a small leak produces. Fitting more freedom to the data is not a remedy with seven independent two-week spans. The data themselves are part of the difficulty: quiet steps are chosen by the absence of a reported repair, and unreported leaks that start or grow inside them inflate the variance in exactly the directions leaks occupy.

The two failures of the validation point in different directions. Detector (i) found slightly more of the 5 m^3/h leaks than the model predicted: in the directions of single-junction leaks the model overstates the noise. The scan's false alarms point at the pump's suction junction, where the pump's flow, which follows the tank level, moves the night heads in a way no demand term represents: there the model understates the noise. Both are properties of the covariance, not of the linearization, which matched EPANET closely and always on the conservative side.

A slow leak is harder than any of these numbers suggest. A leak that grows linearly to its final size over thirteen weeks differs by only about a thirteenth of that size between two consecutive weeks, so a 7 m^3/h ramp presents the step statistic with about half a cubic meter per hour; no gauge set in Table 4 detects that. The leaks of this study are steps.

The design results are relative, and they hold across the covariances tried: the same few Area A gauges carry most of the information, Area B's single gauge adds least, and the first additions fall in the same regions whichever reasonable model of the noise is used, except the one without regional structure. What a demand measurement is worth depends on that structure, and so does how small a leak can be detected. The linearization is taken at one night operating point with the leaks of the day absent; the repairs show that a network carrying several large leaks responds more strongly per unit leak, so the sensitivities, and the smallest detectable leak with them, depend on the state of the network at the time.

6 Conclusion

This report asked four questions of L-Town's 30 Area A and B pressure gauges and their 2018 night readings. The linearized network reproduces EPANET's response to a leak in direction to a cosine of at least 0.955 and underestimates its size by a median 5 %. The night noise is dominated by the area-wide consumption level and by regional deviations from it; a physical model captures its directions but not its size, failing a whitening check in 30 of 30 directions. The existing gauges carry most of their information in a few Area A sites, lose little when their least valuable third is removed, and gain little from ten more (0.172 bits per leak); a known demand scale is worth less than one gauge under the fitted model and much more under a model without regional structure. The validation fixed in advance failed: the noise model does not predict detection on real 2018 pressures; placements are ranked by relative information only, and no statement is made on whether a leak of a given size would be detected. Under the model the median smallest detectable leak is 6.1 m^3/h with today's gauges and 5.4 with ten more, and it spans 4.2 to 6.1 m^3/h across the covariances considered. Where a utility would put its next gauge is answered; how small a leak it would then see is not, until the night noise is modeled to within the size of a leak.

A Reproduction

```
cd docs/terra_graph_engine/case_studies/water_networks/sensor_placement_l_town
PYTHONPATH=../../../../../src python run.py
PYTHONPATH=../../../../../src python make_report.py --pdf
```

`run.py` reads only the study's committed `data/` and the nominal network file, and writes `results.json` (every number at six significant digits) and `timings.json`; `make_report.py` reads only `results.json`. The committed inputs are regenerated by `data/make_excerpt.py` from the verified 2018 files and by `run.py --stage epanet`, which needs EPANET through WNTR. Seeds: 20180918 for the injections, 20180101 for every threshold (20000 draws for the validation, 5000 for the design curves), 20180104 for resampling. Only 2018 data are read.

B Data and problems

The BattLeDIM dataset (Vrachimis et al., 2022; Zenodo 10.5281/zenodo.4017659, CC BY 4.0) supplies the nominal network, the 2018 readings and the 2018 repair report. The study commits the 2018 night means of the 30 gauges and three flow meters, the 03:00 operating point, the EPANET check responses and the injection signatures. Under `problems/`, four directories: *leak head sensitivity on a loop* (the conductance solve by hand), *distinguishability behind a PRV* (why leaks in a valve-fed zone look alike from outside it), *matched SNR with correlated noise* (two gauges with a shared mode) and *a second gauge beside the first* (the diminishing return of a correlated gauge). Each is a statement, a script that solves it with the plugin, and its results file.

References

[Vrachimis et al. 2022] Vrachimis, S. G., Eliades, D. G., Taormina, R., Kapelan, Z., Ostfeld, A., Liu, S., Kyriakou, M., Pavlou, P., Qiu, M., Polycarpou, M. M. (2022). Battle of the Leakage Detection and Isolation Methods. Journal of Water Resources Planning and Management 148(12), 04022068. doi:10.1061/(ASCE)WR.1943-5452.0001601.

[Rossman et al. 2020] Rossman, L. A., Woo, H., Tryby, M., Shang, F., Janke, R., Haxton, T. (2020). EPANET 2.2 User Manual. U.S. Environmental Protection Agency, EPA/600/R-20/133.

[Perez et al. 2011] Pérez, R., Puig, V., Pascual, J., Quevedo, J., Landeros, E., Peralta, A. (2011). Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks. Control Engineering Practice 19(10), 1157–1167. doi:10.1016/j.conengprac.2011.06.004.

[Krause et al. 2008] Krause, A., Singh, A., Guestrin, C. (2008). Near-optimal sensor placements in Gaussian processes: theory, efficient algorithms and empirical studies. Journal of Machine Learning Research 9, 235–284.

[Das and Kempe 2011] Das, A., Kempe, D. (2011). Submodular meets spectral: greedy algorithms for subset selection, sparse approximation and dictionary selection. In: Proceedings of the 28th International Conference on Machine Learning, pp. 1057–1064.

[Wilson 1927] Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association 22(158), 209–212. doi:10.1080/01621459.1927.10502953.