Bayesian inference of state, leak size and leak location in a small water network from two pressure gauges

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

September 2026

Abstract

In a water distribution network the heads and flows follow from conservation at the junctions and the head-loss law of the pipes once the demands are known, and a leak breaks that balance in a way pressure gauges can see. Before such inference is trusted on a city network, its equations should reproduce answers that can be checked by hand. This study does that on the smallest network with a loop and a series element: a reservoir feeding a junction through two parallel pipes and a second junction through a third, with metered demands, two pressure gauges and a possible leak at the second junction. Written as conservation and Hazen–Williams residuals, the same equations answer three questions. Deterministic heads and flows agree with the parallel-pipe closed form to 4.0×10^{-15} relative and with EPANET 2.2 to 3.8×10^{-4} m. With gauges of 10 cm noise and a 10 kg/s leak, the leak-free model is rejected in 20 of 20 noise draws, the leak's size is recovered as 10.04 kg/s with a posterior standard deviation of 0.17 kg/s (18 of 20 95 % intervals cover the truth, against 19 expected), and the leaking junction is identified in 20 of 20 draws with posterior mass at least 1.00. A sweep of smaller leaks locates the detection threshold of these two gauges: 1 kg/s (a 1 % loss) is found in every draw, 0.25 kg/s is not found at all, while the size estimate stays unbiased throughout. Every number costs 218 engine solves.

1 Introduction

Water utilities lose a substantial share of the water they treat to leaks, and locating a leak from the few pressure gauges and flow meters a network carries has been a research problem for decades; Puust, Kapelan, Savic and Koppel (2010) review the methods. Every one of them rests on the same physics: mass is conserved at the junctions, and the head lost along a pipe is a fixed function of its flow, so a leak, an unmetered outflow, changes the pressures the gauges read. Whether a method can be trusted on a city network depends first on whether its equations reproduce answers that can be checked exactly, and small networks with a closed-form solution and an independent reference solver are where that check is possible.

The reference solver for water networks is EPANET (Rossman et al., 2020), which solves the same conservation and head-loss equations for the deterministic state and serves here as the independent check. Writing those equations as a factor graph and inferring the state and its faults from sparse readings is the construction of Han, Eguchi, Mehrotra and Venkatasubramanian (2018) for a network damaged by an earthquake and of Irofti, Romero-Ben, Stoican and Puig (2025) for leak localization over a time window; both rest on the sparse least-squares machinery that Dellaert and Kaess (2017) describe for robotics. Comparing leak hypotheses by their Bayesian evidence, so that the answer is a probability per candidate rather than a score, is older than all three.

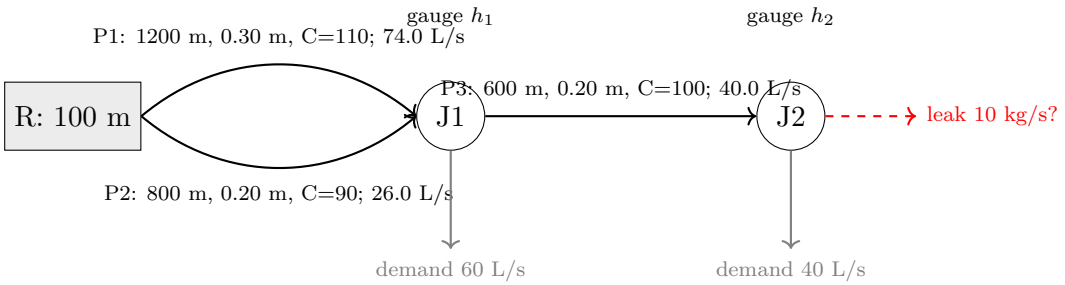


Figure 1: The network of the study: a reservoir, two pipes in parallel, one in series, two metered demands, two pressure gauges, and a possible leak at J2. Flows are the leak-free closed form.

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, attached to the same graph, whatever the sensors report and whatever is known in advance. Each of these is a relation that holds only to within a stated tolerance, and the engine weighs them together to find the most probable state of the system, the uncertainty of every quantity in it, and the relative plausibility of competing explanations of the same readings. Domains enter through plugins that supply their laws, units and file formats; the inference and the bookkeeping are shared. This study uses its hydraulics plugin to write a three-pipe network's mass balances and Hazen–Williams pipe laws, solves the leak-free state against a closed form and EPANET, and ranks three leak hypotheses by how well each explains two gauge readings. Section 3 states the framework and each of these constructs precisely.

This study makes no claim of novelty. On a network small enough to check by hand it asks three questions: what the heads and flows are, and whether the equations reproduce the closed form and EPANET; whether a leak at a known junction can be detected and sized from two gauges, and whether the uncertainty on its size is honest; and whether the leaking junction can be told from its neighbor and from no leak at all, and at what leak size that stops being possible. Section 2 states the network, the data and the model, Section 3 the methods, Section 4 the results, Section 5 their limits, Section 6 the answers, and the appendices the reproduction recipe, the data and the references.

2 System and data

The network. Figure 1 draws the network. The reservoir holds a head of 100 m (the water's energy per unit weight, in meters). Pipes P1 and P2 run in parallel from the reservoir to junction J1; pipe P3 runs from J1 to J2. J1 delivers 60 L/s and J2 40 L/s, both metered and treated as known. The pipes' lengths, diameters and Hazen–Williams roughness coefficients are in the figure; the head-loss coefficients they imply, defined in Section 3, are $K_1 = 747$, $K_2 = 5,207$ and $K_3 = 3,213$ in $\mathrm{s}^{1.852}\,\mathrm{m}^{-4.556}$. Gauges at J1 and J2 read head with a standard deviation of 10 cm.

The data. Everything is synthetic and stated. The truth of each noise draw is the network's own solution with the leak present; the readings are that truth plus Gaussian noise of the gauges' standard deviation, drawn afresh per seed (20 seeds for the main questions, 10 per leak size in the sweep). The demands are known exactly, the reservoir head is fixed, the water density is 1000 $\mathrm{kg/m}^3$, and the leak, when present, is a constant outflow at J2 of 10 kg/s; the sweep repeats the experiment at 0.25, 0.5, 1, 2, 5 kg/s. No field data are used.

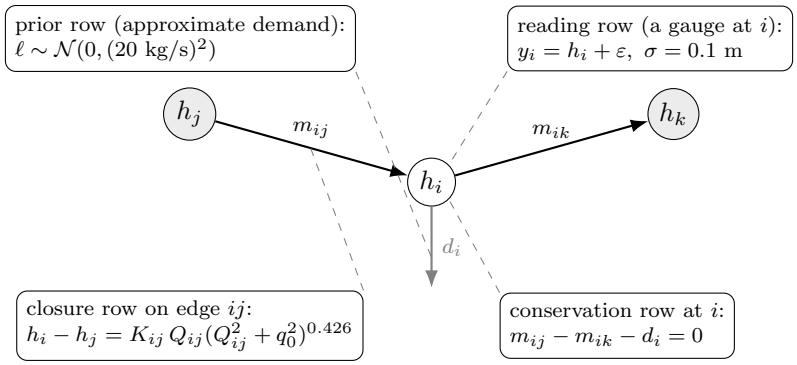


Figure 2: The model around one node. Unknowns are the head at the node, the mass flows on its links and its demand; the equations tying them are conservation at the node, a closure law on each link, a reading row per sensor and a prior row per approximately known quantity. Every row carries a weight; the fit minimizes the weighted squared misfits of all rows at once, and the uncertainty of each unknown follows from the curvature of that misfit. The same picture holds for every node and every domain; only the closure law changes.

The model. Figure 2 shows the equations around one node. Here the whole network has five unknowns, the three pipe flows and the two junction heads, and five physical relations, two conservation balances and three pipe laws, stated in Section 3; the gauges add two reading relations, and a leak hypothesis adds one unknown with a prior.

The questions. (1) What are the heads and flows, and do the equations reproduce the closed form and EPANET? (2) Is there a leak at J2 and how large, with what uncertainty? (3) Is the leak at J1 or at J2, or is there none, and down to what leak size can two gauges tell?

3 Methods

3.1 The three-pipe factor graph

The unknowns are the mass flows m_1, m_2, m_3 (kg/s) on the three pipes and the heads h_1, h_2 (m) at the two junctions; the reservoir head h_R and the demands d_1, d_2 are known. Each physical law is written as a residual, an expression that vanishes when the law holds and carries a weight, the inverse variance of how far it may be off; Section 3.2 makes this the general rule. Conservation at the two junctions and the Hazen–Williams law on the three pipes give five residuals,

$$r_1 = m_1 + m_2 - m_3 - d_1, \tag{1}$$

$$r_2 = m_3 - d_2 - \ell, \tag{2}$$

$$r_3 = h_R - h_1 - K_1 Q_1 (Q_1^2 + q_0^2)^{0.426}, \tag{3}$$

$$r_4 = h_R - h_1 - K_2 Q_2 (Q_2^2 + q_0^2)^{0.426}, \tag{4}$$

$$r_5 = h_1 - h_2 - K_3 Q_3 (Q_3^2 + q_0^2)^{0.426}, \tag{5}$$

$$K_e = \frac{10.667\, L_e}{C_e^{1.852} D_e^{4.871}}, \tag{6}$$

where ℓ is the leak outflow at J2, zero in the leak-free model, $Q_e = m_e/\rho$ the volumetric flow in m^3/s, K_e the head-loss coefficient of a pipe of length L_e, diameter D_e and roughness C_e in SI units,

and q_0 a small regularizing flow that keeps the slope of the law finite at zero flow, at a relative error of $0.426\,(q_0/Q)^2$ against the textbook law. The regularization is set to 1.0×10^{-8} m^3/s for the closed-form comparison and to the plugin's default $q_0 = v_0 A_e$ with v_0 equal to 1.0×10^{-3} m/s everywhere else. The physics residuals are weighted at $(1\text{ g/s})^{-2}$ on conservation and $(1\text{ mm})^{-2}$ on the pipe laws, so the physics is trusted far more than any reading. The two gauges and, in a leaking hypothesis, the prior on the leak attach to the same unknowns as residuals of the same form,

$$r = y_i - h_i, \tag{7}$$

$$r = \ell, \tag{8}$$

weighted by σ_h^{-2} with $\sigma_h = 10$ cm and by τ^{-2} with $\tau = 20$ kg/s, a zero-centered prior wide enough to say nothing about the leak's size.

3.2 The probabilistic model

The residuals of the previous subsection are all of one kind, and the engine treats them as such; this subsection states the general model they belong to. 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, one flow per edge, and whatever else is uncertain, such as demands or parameters. Every relation the state must satisfy is a factor, written as a residual $r_k(z)$ that vanishes when the relation holds and carrying a weight w_k, the inverse variance of how far it may be off. A conservation law at a node, a constitutive law on an edge, a sensor reading and a prior on an approximately known quantity are all factors of this one kind; they differ only in their residual and their weight, a physics relation being trusted to a small tolerance, a reading to its sensor's noise, a prior to its stated spread. Treating each factor as an independent Gaussian belief about its residual gives the posterior density of the state given the readings y,

$$p(z \mid y) \propto \prod_k \exp\!\big(-\tfrac{1}{2}\, w_k\, r_k(z)^2\big). \tag{9}$$

The state that maximizes it, the *maximum-a-posteriori* (MAP) state z^*, minimizes the weighted sum of squared residuals $C(z)$ and is found by Gauss–Newton iteration on the sparse Jacobian J of the residuals, $J_{ki} = \partial r_k / \partial z_i$, with W the diagonal matrix of weights, from a flat start until the gradient norm falls below a tolerance ϵ:

$$C(z) = \tfrac{1}{2} \sum_k w_k\, r_k(z)^2, \tag{10}$$

$$z^* = \arg\min_z C(z), \tag{11}$$

$$z \leftarrow z - (J^\top W J)^{-1} J^\top W\, r(z) \quad \text{until } \|J^\top W\, r(z)\| < \epsilon. \tag{12}$$

Near the minimum the cost (10) is approximately quadratic, so the posterior is approximately Gaussian about z^* with information matrix Λ. This is the *Laplace approximation*, and the *posterior standard deviation* σ_i of any unknown is read from the diagonal of Λ^{-1}, from the same sparse factorization the last Gauss–Newton step used:

$$p(z \mid y) \approx \mathcal{N}\big(z^*, \Lambda^{-1}\big), \tag{13}$$

$$\Lambda = J^\top W J, \tag{14}$$

$$\sigma_i = \sqrt{(\Lambda^{-1})_{ii}}. \tag{15}$$

When several factor graphs compete to explain the same readings — a leak here, a leak there, no leak — each is compared by its *evidence* $Z = \int \prod_k \exp(-\frac{1}{2} w_k r_k(z)^2)\, dz$, the probability of the readings under the model once the extra freedom of its unknowns has been paid for. In the Laplace approximation the evidence has a closed form, and normalizing every factor by its own weight lets graphs with different numbers of unknowns d and factors compare as densities,

$$\log Z = -C(z^*) + \tfrac{d}{2} \log 2\pi - \tfrac{1}{2} \log\det \Lambda + \sum_k \tfrac{1}{2} \log \frac{w_k}{2\pi}. \tag{16}$$

Differences of $\log Z$ between hypotheses, from (16), are in *nats*, natural-log units of odds; five nats is odds of about 150 to 1. The posterior probability of a hypothesis is its evidence normalized over all hypotheses under equal prior odds. Everything that follows is a choice of variables and factors on this graph, followed by these three operations.

3.3 Reference solutions

Two independent answers exist for the leak-free state. Two pipes in parallel see the same head drop, so the flows split in a fixed ratio and the closed form follows,

$$K_1 Q_1^{1.852} = K_2 Q_2^{1.852}, \tag{17}$$

$$Q_1/Q_2 = (K_2/K_1)^{1/1.852}, \tag{18}$$

$$Q_1 + Q_2 = d_1 + d_2 + \ell, \tag{19}$$

from which the flows and then the heads follow; here $Q_1/Q_2 = 2.8524$ against a predicted 2.8524 from (18). The second answer is EPANET 2.2 (Rossman et al., 2020), the reference hydraulic solver, given the same network as an input file in SI units and run through WNTR 1.5.0 at accuracy 10^{-8}. Against these, with the demands known and no gauges, the five residuals (1)–(5) are a square system and Newton's method solves them exactly from a flat start (heads at the reservoir head, flows at 1 kg/s); this is the iteration (12) with every weight equal.

3.4 Leak hypotheses and their comparison

Three hypotheses compete to explain the two gauge readings: no leak, M_0, in which ℓ is absent; a leak at J1, M_1, in which the leak outflow enters the conservation residual of J1; and a leak at J2, M_2, in which it enters (2) as written. Each leaking hypothesis adds the prior (8). Every hypothesis is solved to its MAP state (11) and its Laplace posterior (13); the leak's 95 % interval is ± 1.96 posterior standard deviations (15). Whether a hypothesis fits at all is tested by the chi-square misfit of its reading residuals at the optimum against the stated noise,

$$\chi^2 = \sum_{k \in \text{readings}} w_k\, r_k(z^*)^2, \tag{20}$$

on as many degrees of freedom as readings minus fitted parameters (two for M_0), rejected at $\alpha = 0.01$. The hypotheses are then compared by their normalized evidence (16), and their posterior probabilities are the evidences normalized under equal prior odds,

$$P(M_j \mid y) = \frac{Z_j}{\sum_k Z_k}. \tag{21}$$

Table 1: Deterministic state without a leak. The Newton solve at $q_0 = 10^{-8}$ m^3/s reproduces the closed form to 4.0×10^{-15} relative; at the plugin's default regularization the heads move by 4.9×10^{-6} m; EPANET 2.2 differs from the closed form by 3.8×10^{-4} m in head and 5.4×10^{-9} m^3/s in flow.

quantity	closed form	Newton, $q_0 = 10^{-8}$	Newton, default q_0	EPANET 2.2
Q_1 [L/s]	74.0422	74.0422	74.0422	74.0422
Q_2 [L/s]	25.9578	25.9578	25.9578	25.9578
Q_3 [L/s]	40.0000	40.0000	40.0000	40.0000
h_1 [m]	93.97740	93.97740	93.97740	93.97760
h_2 [m]	85.70010	85.70010	85.70010	85.70040

3.5 The detection sweep

To find where two gauges stop seeing a leak, the three-hypothesis comparison is repeated with the true leak at J2 set to 0.25, 0.5, 1, 2, 5 kg/s in turn, 10 noise draws each, at the same gauges and noise. For every leak size the run records how often J2 is top, the mean posterior on J2 and on no leak, how often the leak-free model is rejected, and the MAP leak size with its posterior standard deviation.

3.6 Metrics

The questions of Section 2 are answered with measures chosen so that each can be read against a reference. Agreement of the deterministic state is the largest relative difference between the Newton solution and the closed form, whose reference is machine precision, and the largest absolute difference in head and flow against EPANET, whose reference is EPANET's own output precision of about 10^{-4} m. Detection is the number of draws in which the leak-free model is rejected by (20) at $\alpha = 0.01$; under no leak the expected rejection rate would be α, so a leak that is detected should be rejected in nearly every draw. Estimation is the mean MAP leak size against the true 10 kg/s, the spread of the estimate across draws against the mean posterior standard deviation, which a calibrated posterior makes equal, and the number of 95 % intervals that cover the truth, about 19 of 20 for a calibrated interval. Location is the number of draws in which J2 carries the largest posterior probability (21) and the smallest posterior mass it receives; the reference is every draw and a mass near one. The detection threshold is the leak size at which J2 is top in every draw of the sweep, read against the size at which the leak-free model and J2 share the mass. Cost is counted in engine solves, one Newton solve per deterministic answer and one MAP solve per hypothesis per draw, rather than timed, so that it does not depend on the machine.

4 Results

Deterministic state. Table 1 lists the leak-free state four ways. The Newton solve at $q_0 = 10^{-8}$ m^3/s reproduces the closed form to 4.0×10^{-15} relative, which is machine precision; the plugin's default regularization moves the heads by 4.9×10^{-6} m, the stated $0.426\,(q_0/Q)^2$ effect. EPANET differs from the closed form by 3.8×10^{-4} m in head, EPANET's own output precision, so the two independent references agree with the equations to the precision each can offer. A 10 kg/s leak at J2 lowers h_1 by 1.16 m and h_2 by 5.40 m: the signature the gauges must see, 54 times their noise at J2.

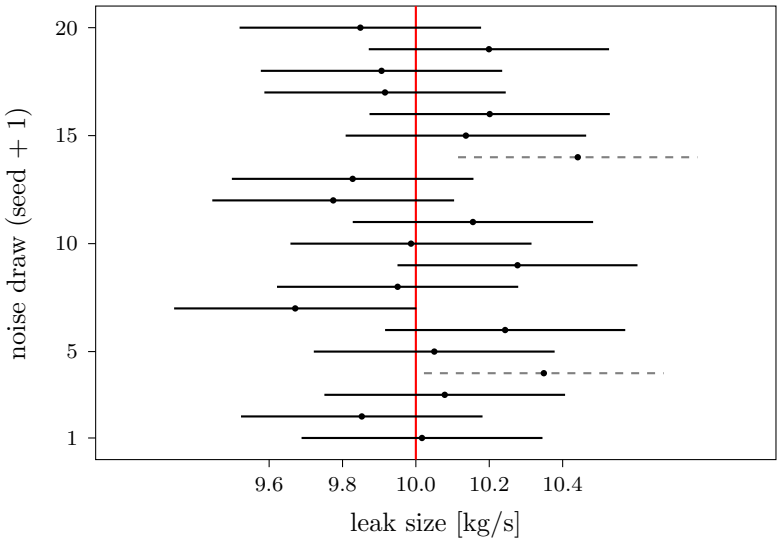


Figure 3: The leak's 95 % Laplace interval for each of the 20 noise draws (dots: MAP estimate; red line: the true 10 kg/s). 18 of 20 intervals cover the truth; the two that miss are drawn dashed.

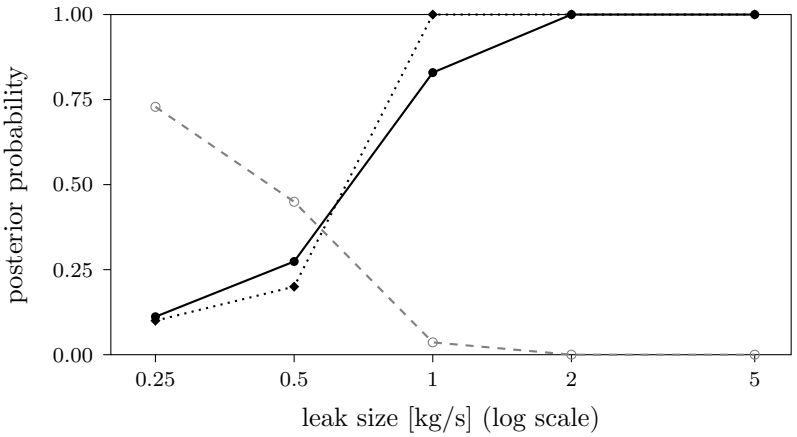
Table 2: Leak location, first noise draw: readings $h_1 = 92.827$ m, $h_2 = 80.288$ m. Normalized Laplace log-evidence and posterior probability under equal prior odds. The leak-free model's misfit is $\chi^2 = 3,060.3$ on 2 degrees of freedom (inadequate); the true-node model's is $\chi^2 = 0.27$ (adequate).

hypothesis	$\log Z$	P
no leak	-1,526.8	0.000
leak at J1	-455.4	0.000
leak at J2	-1.7	1.000

Leak size. With the leak present and 20 independent noise draws on the two gauges, the leak-free model is rejected by (20) at $\alpha = 0.01$ in 20 of 20 draws and the true-node model accepted in 20, so detection is never in doubt at this size. The MAP leak size averages 10.044 kg/s across draws (spread 0.205 kg/s) against a true 10 kg/s, with a mean posterior standard deviation of 0.167 kg/s; the spread across draws and the posterior width agree, which is what a calibrated Laplace posterior should show on a nearly linear problem. Figure 3 draws the 20 intervals: 18 cover the truth, against about 19 expected for a 95 % interval.

Leak location. Table 2 shows the first draw. The leaking node is J2 in 20 of 20 draws; the posterior mass (21) on J2 is never below 1.000, the evidence (16) for J2 exceeds the null by at least 1,413 nats and the J1 alternative by at least 420 nats. The two gauges separate the nodes because a leak at J1 lowers both readings by the same parallel-pipe drop while a leak at J2 also adds P3's drop to h_2.

Detection threshold. Figure 4 and Table 3 repeat the three-hypothesis comparison for smaller leaks at the same gauges. At 1 kg/s — a signature of 50 cm at J2 against 10 cm of noise — J2 is on top in every draw with mean posterior 0.83; at 0.5 kg/s the null and the true node share the mass (0.45 against 0.27); at 0.25 kg/s the null wins with 0.73. The MAP size stays unbiased throughout (0.31 kg/s for a 0.25 kg/s leak) with a posterior standard deviation near 0.20 kg/s set



solid: mean P(leak at J2); dashed: mean P(no leak); dotted: fraction of draws with J2 on top

Figure 4: Detection sweep at the same two gauges (10 cm noise), 10 noise draws per leak size. The posterior mass on the true node rises from 0.11 at 0.25 kg/s to 0.83 at 1 kg/s and 1.00 from 2 kg/s.

Table 3: Detection sweep. Signature: the head drop the leak causes at J2. Columns: draws with J2 on top, mean posterior on the true node and on the null, draws in which the leak-free model is rejected at $\alpha = 0.01$, and the MAP leak size with its mean posterior standard deviation.

leak [kg/s]	signature [cm]	J2 on top	$\bar{P}$(J2)	$\bar{P}$(none)	null rejected	size [kg/s]
0.25	12.4	1/10	0.11	0.73	2/10	0.31 ± 0.20
0.5	24.9	2/10	0.27	0.45	3/10	0.56 ± 0.20
1	49.9	10/10	0.83	0.04	10/10	1.06 ± 0.19
2	100.8	10/10	1.00	0.00	10/10	2.05 ± 0.19
5	258.7	10/10	1.00	0.00	10/10	5.05 ± 0.18

by the gauges, so the threshold is a matter of evidence, not of estimation: below about 0.5 kg/s the extra parameter is not worth what (16) charges for it.

Cost. 8 Newton solves and 210 MAP solves (218 in all) produce every number in this report.

5 Discussion

Everything here is synthetic and small by design: the truth is the network's own solution, the noise is Gaussian at a stated 10 cm, and the demands are known exactly, which no field network offers. Those choices are what let the closed form and EPANET serve as independent checks of the equations, and they are also the limits of the result: with uncertain demands a leak competes with demand at the same node and the detection threshold moves up; with fewer or noisier gauges the three hypotheses blur sooner than the sweep shows. A single steady snapshot cannot use the constancy of a leak across a day, which is the construction that separates a leak from a varying demand on larger networks. The comparison with the cited works is one of formulation, not of numbers: they were not run here, and this network is too small to discriminate methods.

6 Conclusion

This report asked three questions of one set of conservation and Hazen–Williams residuals on a three-pipe network watched by two pressure gauges. Whether the equations reproduce the known state: they do, to 4.0×10^{-15} relative against the closed form and 3.8×10^{-4} m against EPANET 2.2. Whether a 10 kg/s leak at the second junction can be detected and sized, with honest uncertainty: the leak-free model is rejected in 20 of 20 draws, the size is recovered as 10.04 kg/s with a posterior standard deviation of 0.17 kg/s, and 18 of 20 intervals cover the truth against about 19 expected. Whether the leaking junction can be told from its neighbor and from no leak, and down to what size: J2 carries the largest probability in all 20 draws with at least 1.00 of the mass, and in the sweep the threshold of these two gauges lies near 0.5 kg/s, where the leak-free model and the true node share the mass; 1 kg/s is found in every draw and 0.25 kg/s in none, while the size estimate stays unbiased throughout, so the threshold is one of evidence rather than estimation. None of the constructions is new; the study fixes the form in which the larger studies that follow report and leaves four exercise-sized problems in the corpus. The same equations are next applied to a city-scale network with uncertain demands and to leak localization over a day of readings.

A Reproduction

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

`run.py` writes `results.json` (every number at six significant digits) and `timings.json`; `make_report.py` reads only `results.json`. Seeds 0–19 drive the gauge noise; the sweep uses seeds 0–9. EPANET 2.2 is run through WNTR 1.5.0 (the `oracles` extra). Water density 1000 kg/m^3.

B Data and problems

The network is the study's own construction (no external data). Under `problems/`, four directories: *parallel split* (the split ratio and heads, closed form against Newton); *pump–system intersection* (a one-point EPANET pump curve against a system curve, scalar root against Newton); *leak from two readings* (misfit of the leak-free model and the MAP leak size with its interval); *which node* (three hypotheses on the normalized evidence). Each is a statement, a script that solves it with the plugin, and its results file.

References

[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.

[Han et al. 2018] Han, Q., Eguchi, R. T., Mehrotra, S., Venkatasubramanian, N. (2018). Enabling state estimation for fault identification in water distribution systems under large disasters. In: 2018 IEEE 37th Symposium on Reliable Distributed Systems (SRDS), pp. 161–170. doi:10.1109/SRDS.2018.00027.

[Irofti et al. 2025] Irofti, P., Romero-Ben, L., Stoican, F., Puig, V. (2025). Factor graph optimization for leak localization in water distribution networks. arXiv:2509.10982. doi:10.48550/arXiv.2509.10982.

[Puust et al. 2010] Puust, R., Kapelan, Z., Savic, D. A., Koppel, T. (2010). A review of methods for leakage management in pipe networks. Urban Water Journal 7(1), 25–45. doi:10.1080/15730621003610878.

[Dellaert and Kaess 2017] Dellaert, F., Kaess, M. (2017). Factor graphs for robot perception. Foundations and Trends in Robotics 6(1–2), 1–139. doi:10.1561/2300000043.