Calibrating a district heating network from operating data: a posterior over pipe resistances and valve characteristics, and what each set of sensors identifies

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

September 2026

Abstract

A hydraulic model of a heating network is only as good as its parameters, and the resistances of pipes and the characteristics of control valves are rarely known from drawings; they must be inferred from the pressures and flows the network's few sensors record. This study infers them for a laboratory district heating line of four pipe sections and four valve-controlled consumers, recorded by Agner et al. (2024) under 2001 excitation and 1341 realistic load conditions. The line is written as one factor graph over the load conditions, with the four section resistances and the flow coefficient of each valve at 8 positions as shared unknowns; a maximum-a-posteriori fit to the pump's pressure rise and the four consumer flow meters over 1259 training conditions, with the flow-meter noise calibrated on the residuals, gives every parameter with a standard deviation. The fit predicts the flows of 540 held-out excitation conditions to 0.9–1.7 % of the mean flow, against 1.8–3.5 % for the best of the three published models refit by their own estimator, and the flows of the realistic set to 3.1–5.6 % against 3.3–5.4 %. Its 95 % predictive intervals cover 96 % of the held-out excitation flows but only 58 % of the realistic ones, whose error is systematic. Three sensor sets are compared: the plant's sensors determine the first three sections' resistances to 1 % or better but the last one only to 12 %, along a direction that trades it against the last valve's curve; the consumer differential-pressure sensors narrow it 60-fold, at the price of zero offsets of about 0.4 mWC that the fit must absorb and a worse flow prediction; the flow meters alone fix only ratios: the first section's resistance drops out of every equation and sits at its prior, and the scale of the rest is the priors' choice.

1 Introduction

District heating networks, which Werner (2017) reviews, are operated from a few sensors at the plant and at some of the substations, and the decisions taken on them, how hard to pump, how to balance the branches, whether a consumer is drawing what it should, rest on a hydraulic model of the network. That model's parameters are not known from drawings: pipe resistances change with deposits and roughness, control valves follow characteristics that manufacturers specify only nominally and that wear alters, and a model whose parameters are wrong predicts flows that are wrong in proportion. Calibrating the model from operating data is therefore the first task of any model-based operation, and it has a long history in water distribution, where Savic, Kapelan and Jonkergouw (2009) review two decades of methods and their difficulties: too few sensors, parameters that trade off against one another, and results whose uncertainty is rarely stated.

Agner, Møller Jensen, Rantzer, Skovmose Kallesøe and Wisniewski (2024) brought this problem into the laboratory. On the Smart Water Infrastructure laboratory's line network in Aalborg they

recorded two experiments, one that drives the four consumers' valves through their operating region and one that follows realistic load curves, and fitted pipe resistances and parameterized valve curves by regularized least squares, reaching flow predictions within 5 to 10 % of the mean flow; they published the data and the code. What their method does not give is the uncertainty of the parameters, the interval a prediction deserves, or a statement of which parameters the available sensors can and cannot determine. Writing a network's conservation laws as a factor graph, a set of weighted residuals over its unknowns, and inferring the unknowns from sparse readings is the construction of Han, Eguchi, Mehrotra and Venkatasubramanian (2018) for water networks; the sparse least-squares machinery behind it is that of Dellaert and Kaess (2017). Nothing in this report is new; its purpose is to apply that construction to Agner et al.'s data and to answer the three questions their method leaves open.

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 laboratory heating line as one factor graph over every recorded load condition, with the pipe resistances and the valve characteristics as shared unknowns, fits them to the plant's sensors, predicts the flows of held-out conditions with intervals, and reads off what each set of sensors identifies. Section 3 states the framework and each of these constructs precisely.

This report asks three questions of the laboratory line. What are the section resistances and the valve characteristics, and how well does the excitation experiment determine them? How well does the fitted model predict the flows of conditions it has not seen, and are its predictive intervals honest? And what does each set of sensors identify, the plant's, the plant's with the consumers' pressure sensors, and the flow meters alone? Section 2 states the line, 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 line. A pumping station holds a pressure rise dp_{pump} between the return and the supply header; four pipe sections in series carry the water to four consumer modules, each a control valve at a commanded position v_i between 0 and 1, and four return sections carry it back. The pipe unit's sections are physically identical on the supply and return sides, so one resistance s_p per section stands for both. Pressures are recorded in meters of water column (mWC) and flows in liters per minute, and the study keeps those units: resistances are then mWC per $(\mathrm{L/min})^2$ and flow coefficients L/min per $\sqrt{\mathrm{mWC}}$, and every parameter sits within a few decades of one.

The data. Agner et al. (2024) recorded the line at 10 Hz in two experiments and averaged the logs into load conditions, a procedure reproduced here step for step. The excitation set drives the valves in steps over positions 0.3 to 1: 2001 conditions, of which 202 with a consumer below 0.1 L/min (a shut valve, whose coefficient is zero and which carries no information about its curve) are excluded, and the first 70 % of the remaining 1799 (1259 conditions) train the model while

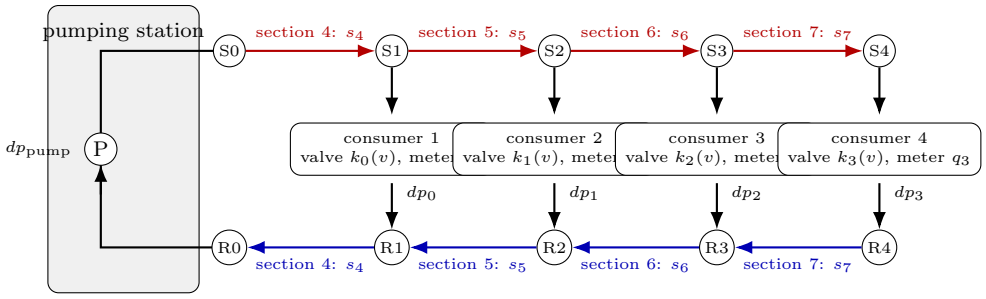


Figure 1: The laboratory line of Agner et al. (2024): a pumping station drives water through four pipe sections to four consumer modules, each a control valve whose position is commanded, and back through four return sections. The unknowns of this study are the resistance s_p of each section (supply and return together) and the flow coefficient $k_i(v)$ of each valve as a function of its position. The sensors are the pump's pressure rise, one flow meter per consumer and one differential-pressure sensor per consumer connection.

the last 540 are held out, the same split the published models use. The realistic set follows load curves with set-point control: 1341 conditions at lower flows and lower valve openings than most of the excitation set. Each condition records the four consumer flows q_i, the four valve positions, the pump's pressure rise and the differential pressure dp_i across each consumer's connection. The valves show hysteresis; Agner et al.'s compensation, which lets the effective position follow the commanded one only once it has moved by more than 1.5 %, is applied to the commanded positions before any model is built and is inherited as is.

The model. Figure 2 shows the equations around one header. Each load condition is a copy of the line with its own heads and flows; the 36 parameters, four section resistances and 8 flow-coefficient values per valve, are shared by every copy, and it is through them that 1259 conditions inform one another; the priors (9) and (10) hold every parameter only loosely. The plant sensor set reads the pump's pressure rise and the four flows of every training condition (27698 rows in all); the second set adds the four consumer differential pressures, each with a zero offset as a further unknown; the third reads the flows alone.

The questions. (1) What are the resistances and the valve characteristics, and how well are they known? (2) How well does the model predict unseen flows, and are its intervals honest? (3) What does each sensor set identify, and along which direction does it fail?

3 Methods

3.1 The line as a factor graph over load conditions

For load condition c the unknowns are the heads $h_{n,c}$ at the eight headers other than the return header, which holds the reference, the flows $q_{e,c}$ on the pump, the eight pipe sections and the four consumer branches, and the pump's lift H_c; the parameters shared by every condition are the section resistances $s_4, \dots, s_7$ and the flow-coefficient values k_i^j of valve i at knot j. Each physical law is a residual, an expression that vanishes when the law holds and carries a weight, the inverse

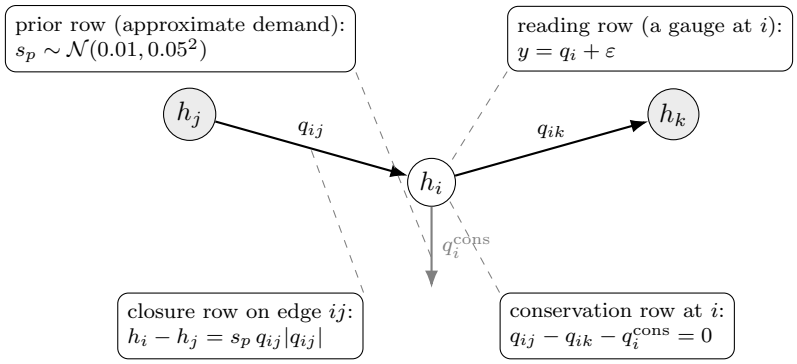


Figure 2: The model around one header of the line, repeated for every load condition. Unknowns are the head at the header, the flows on its pipe sections and into its consumer, and, shared by every load condition, the resistance of each section and the flow coefficient of each valve at fixed positions; the equations are a mass balance at the header, a resistance law on each section and a valve law on each consumer branch, a reading row per sensor and a prior row per parameter. Every row carries a weight; the fit minimizes the weighted squared misfits of all rows over all load conditions at once, and the uncertainty of each unknown follows from the curvature of that misfit.

variance of how far it may be off; Section 3.2 makes this the general rule. Mass is conserved at every header, the head drop along a pipe section follows a quadratic resistance law, the pump imposes its lift, and a consumer branch drops the whole head between its supply and return header through its valve,

$$r^m_{n,c} = \sum_{e \to n} q_{e,c} - \sum_{e \leftarrow n} q_{e,c}, \tag{1}$$

$$r^h_{e,c} = h_{i,c} - h_{j,c} - s_p\, q_{e,c}|q_{e,c}|, \tag{2}$$

$$r^H_c = h_{S0,c} - h_{R0,c} - H_c, \tag{3}$$

$$r^v_{i,c} = h_{S,c} - h_{R,c} - \frac{q_{i,c}|q_{i,c}|}{k_i(v_{i,c})^2}, \tag{4}$$

$$k_i(v) = \sum_j w_j(v)\, k_i^j, \tag{5}$$

where $k_i(v)$, the valve's flow coefficient (flow per square root of head), is the piecewise-linear interpolation (5) of its values at the knots $v \in \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$ with the hat weights $w_j(v)$, which sum to one and are non-zero only at the two knots bracketing v. The quadratic law with exponent 2 is Agner et al.'s choice and is confirmed on these data at full opening, where hysteresis is absent. The four consumer pipes of the laboratory are folded into their valves' curves, since a constant added to a branch's resistance cannot be told from the curve; the published model carries them as separate resistances that its estimator sets to zero. The physics residuals are weighted at $(10^{-3})^{-2}$, so the physics is trusted far more than any reading. The readings and the

priors attach to the same unknowns as residuals of the same form,

$$r = y_c^{dp} - H_c, \tag{6}$$

$$r = y_{i,c}^{q} - q_{i,c}, \tag{7}$$

$$r = y_{i,c}^{dp_i} - (h_{S,c} - h_{R,c}) - b_i, \tag{8}$$

$$r = s_p - \bar{s}, \tag{9}$$

$$r = k_i^j - \bar{k}, \tag{10}$$

weighted by σ_p^{-2} with $\sigma_p = 0.02$ mWC on the pressure readings, by $\sigma_{q,i}^{-2}$ on the flow readings with $\sigma_{q,i}$ calibrated below, and by the priors' inverse variances, $\bar{s} = 0.01$ with standard deviation 0.05 mWC per $(\mathrm{L/min})^2$ and $\bar{k} = 2$ with standard deviation 3 L/min per $\sqrt{\mathrm{mWC}}$, both wide enough to say almost nothing. In the second sensor set the consumer pressure reading (8) carries the sensor's zero offset b_i as an unknown with a zero-mean prior of standard deviation 0.5 mWC. Only the training conditions enter the fit; the held-out and realistic conditions are predicted from it as described below.

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,

$$p(z \mid y) \propto \prod_k \exp\!\big(-\tfrac{1}{2}\, w_k\, r_k(z)^2\big). \tag{11}$$

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:

$$C(z) = \tfrac{1}{2} \sum_k w_k\, r_k(z)^2, \tag{12}$$

$$z^* = \arg\min_z C(z), \tag{13}$$

$$z \leftarrow z - (J^\top W J)^{-1} J^\top W\, r(z) \quad \text{until } \|J^\top W\, r(z)\| < \epsilon. \tag{14}$$

Near the minimum the cost (12) 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:

$$p(z \mid y) \approx \mathcal{N}\big(z^*, \Lambda^{-1}\big), \tag{15}$$

$$\Lambda = J^\top W J, \tag{16}$$

$$\sigma_i = \sqrt{(\Lambda^{-1})_{ii}}. \tag{17}$$

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,

$$\log Z = -C(z^*) + \tfrac{d}{2} \log 2\pi - \tfrac{1}{2} \log\det \Lambda + \sum_k \tfrac{1}{2} \log \frac{w_k}{2\pi}. \tag{18}$$

Differences of $\log Z$ between hypotheses, from (18), 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 Calibrating the flow-meter noise

The flow meters' noise is not documented, and it sets both the weight of every flow reading and the width of every predictive interval. It is taken from the data in two passes: a first fit at a nominal $\sigma_q = 0.15$ L/min, then, with the root-mean-square residual of each consumer's training flows,

$$\sigma_{q,i} = \sqrt{\frac{1}{N_{\text{train}}} \sum_c \left(y^q_{i,c} - q^*_{i,c} \right)^2}, \tag{19}$$

a second fit at those values, here 0.044, 0.036, 0.035, 0.036 L/min for consumers 1 to 4. This makes the training misfit consistent with the stated noise by construction, so the test of the noise model is the held-out set, not the training set.

3.4 Predictions and their intervals

For a held-out or realistic condition the flows are solved forward from the pump's pressure rise and the valve positions at the MAP parameters θ^*, and the parameters' uncertainty is carried through to first order: with $g_{i,c} = \partial q_{i,c} / \partial \theta$ the sensitivity of the flow to the parameters, taken by finite differences of that forward solve, and Σ_θ the parameters' posterior covariance from (15), the predictive variance adds the flow-meter noise,

$$\sigma^2_{\text{pred},i,c} = g_{i,c}^\top \Sigma_\theta\, g_{i,c} + \sigma^2_{q,i}, \tag{20}$$

and the 95 % interval is $\pm 1.96\, \sigma_{\text{pred}}$ around the forward solution. This is the same Laplace linearization the posterior itself rests on; its coverage, the fraction of measured flows inside the interval, tests whether the posterior is honest.

3.5 Identifiability under each sensor set

For each sensor set the marginal posterior of the parameters is compared with the prior: a parameter whose posterior standard deviation is below a quarter of its prior's is called well identified, above three quarters not identified, and weakly in between. The parameters' marginal information matrix, the Laplace information (15) with every other unknown eliminated by Schur complement, has as its flattest eigenvector the direction the sensors constrain least; it is named.

3.6 The reference fit

Agner et al. (2024)'s estimator is re-run on the same training conditions. For every condition and every consumer it writes the pressure balance along the path from the pump through that consumer,

$$dp_{\text{pump}} = \sum_{p \in \text{path}(i)} 2 s_p q_p^2 + q_i^2 \, \theta_i \cdot f(v_i), \tag{21}$$

with f a fixed basis of valve curves, the single function $1/v^2$ in their model A and sixty ramp functions in models B and C, and minimizes the mean absolute residual over non-negative parameters; model C adds the hysteresis compensation used here. The paper solved this with a convex-optimization library; here it is the same linear program solved by SciPy's HiGHS. Flows are then predicted from the pump's pressure rise and the valve positions by the paper's own recursion over equivalent resistances. Where two fits disagree, the source is one of five, input, formulation, scenario, solver or aggregation; here the input and the scenario are the same, and the formulation differs in the valve law and in the estimator.

3.7 Metrics

Parameters are reported as posterior mean and standard deviation, and the valve curves are read against an independent estimate, the median of $q_i/\sqrt{dp_i}$ in bins of the position on the held-out conditions, which uses the consumer pressure sensors the plant fit does not. Prediction is the root-mean-square error of the predicted flow relative to each consumer's mean flow, read against the reference models on the same conditions and against Agner et al.'s published 5 to 10 %. Calibration is the error against the mean predictive standard deviation, equal for an honest posterior, and the fraction of 95 % intervals that cover, about 95 % for an honest one. Identifiability is the count of parameters in each class and the flattest direction. Cost is counted in MAP solves: 8 for everything in this report, subset included.

4 Results

Parameters. Table 1 lists the section resistances and the valves' coefficients at full opening under the three sensor sets, and Figure 3 draws the four valve characteristics inferred from the plant's sensors. The first three sections are known to 1 % or better from the plant's sensors; the fourth is the least determined parameter, 58.27 ± 7.04, a relative standard deviation of 12 %, because the water that passes it passes only the fourth valve, and a resistance in series with a valve whose curve is free is nearly one more constant the curve can absorb; only the curvature of the valve law across the positions the data visit separates the two. The valve characteristics rise nearly linearly from about 0.5 at $v = 0.4$ to about 3.9 at $v = 0.9$ and flatten at full opening; the fourth valve's upper knots carry the widest bands for the reason just given, its curve and the section's resistance being

Table 1: Posterior mean ± standard deviation of the section resistances (in 10^{-3} mWC per $(\mathrm{L/min})^2$, supply and return together) and of each valve's flow coefficient at full opening (L/min per $\sqrt{\mathrm{mWC}}$), under the three sensor sets, from 1259 training conditions. In the last column the first section's resistance sits at its prior mean of 10 and the other parameters are determined only relative to it: without a pressure reading the flows fix only ratios. The consumer pressure sensors' zero offsets, inferred alongside, are in the last row.

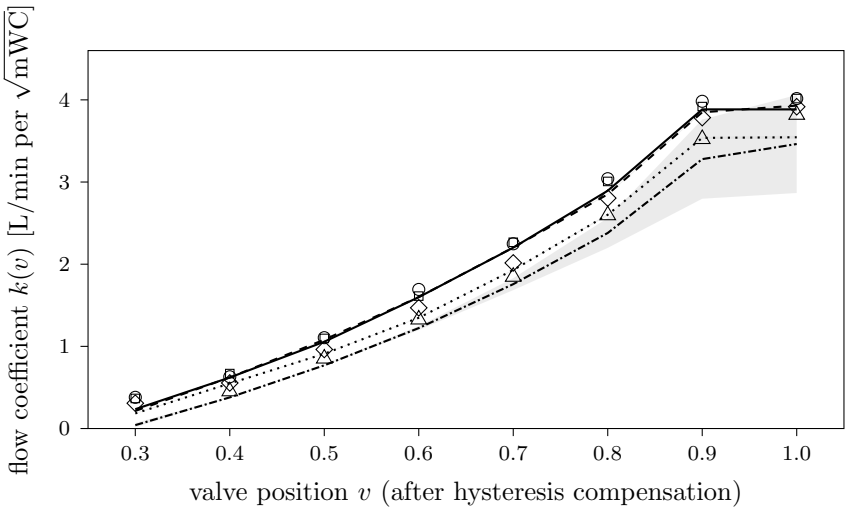
parameter	plant sensors	plant + consumer sensors	flow meters alone
s_4	3.66 ± 0.01	3.40 ± 0.01	10.00 ± 50.00
s_5	4.49 ± 0.03	4.34 ± 0.01	0.05 ± 0.00
s_6	29.63 ± 0.10	31.27 ± 0.04	0.34 ± 0.01
s_7	58.27 ± 7.04	69.24 ± 0.12	-3.00 ± 0.76
$k_0(1.0)$	3.88 ± 0.01	3.82 ± 0.00	35.82 ± 0.68
$k_1(1.0)$	3.93 ± 0.01	3.82 ± 0.00	36.39 ± 0.69
$k_2(1.0)$	3.54 ± 0.01	3.52 ± 0.01	32.57 ± 0.65
$k_3(1.0)$	3.46 ± 0.30	3.59 ± 0.01	10.97 ± 1.00
sensor offsets $b_0..b_3$ [mWC]	—	-0.45, -0.41, -0.30, -0.12	—

Table 2: Root-mean-square error of the predicted consumer flow relative to the mean flow of each consumer (1–4), on the 540 held-out excitation conditions and on the 1341 realistic conditions. The reference rows are Agner et al. (2024)'s three models refit by their own estimator on the same training split and evaluated on the same conditions; model C is their best. Every fit predicts from the pump's pressure rise and the valve positions only.

fit	excitation, held-out [%]				realistic set [%]			
this study, plant sensors (plant sensors)	0.9	0.9	1.1	1.7	3.1	4.0	5.6	3.6
this study, plant + consumer sensors	1.2	1.7	3.1	6.9	3.2	4.7	12.4	16.6
Agner et al., model A (linear valve law)	30.1	27.4	24.7	21.6	34.5	67.0	92.5	34.1
Agner et al., model B (ramp basis)	4.5	4.4	4.4	3.8	5.2	9.4	11.2	8.8
Agner et al., model C (ramp basis, hysteresis)	3.5	2.6	3.5	1.8	3.3	4.1	5.4	3.7

determined largely jointly. Where the consumer pressure sensors allow an independent reading of the same curves, the two agree within 41 % over 31 position bins, the largest differences at wide openings, where the differential pressure across a consumer is smallest and a sensor offset of the size found below (Table 1) weighs most on $q_i/\sqrt{dp_i}$; the check is quantitative at low openings and qualitative at high ones.

Predictions. Table 2 compares the flows predicted from the pump's pressure rise and the valve positions with the measurements. On the 540 held-out excitation conditions the plant-sensor fit predicts the four consumers' flows to 0.9–1.7 % of their mean flow, against 1.8–3.5 % for the best published model refit on the same split and 21.6–30.1 % for the linear valve law; the free valve curve of (5) does the work, since the same physics with a single-parameter valve law (model A) is off by a quarter of the flow. On the realistic set the plant fit predicts to 3.1–5.6 % against 3.3–5.4 % for model C, so the two formulations extrapolate about equally well to the regime the excitation set visits least. Table 3 tests the intervals: on the held-out excitation conditions the error and the predictive standard deviation agree and 96 % of the 95 % intervals cover, so the posterior is honest where the data resemble the training data; on the realistic set the error is two to three times the interval and only 58 % cover. Figure 4 shows why: the realistic set moves the valves slowly through



lines: MAP $k(v)$ of valves 1–4 (solid, dashed, dotted, dash-dotted) from the plant's sensors; bands ±2 posterior sd
markers: median $q_i/\sqrt{dp_i}$ per position bin of the held-out conditions (consumer sensors, unused by the fit)

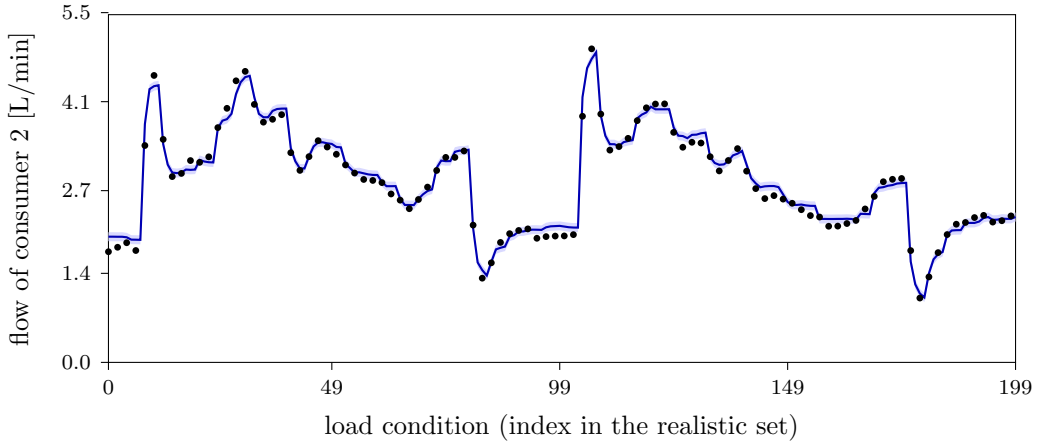
Figure 3: The four valve characteristics inferred from the pump's pressure rise and the flow meters alone, with their ±2 posterior standard deviation bands, against an independent reading of the same curves from the consumer differential-pressure sensors on the held-out excitation conditions. The valves open nearly linearly between 0.3 and 0.9 and saturate at full opening; the fourth valve's curve is the least certain because the plant's sensors separate it only weakly from the last pipe section (Section 4). Over the 31 bins with a coefficient above 0.2 the MAP curve and the binned medians differ by at most 41 %, the most at wide openings where the consumer sensors' offsets weigh most.

low openings, and there the measured flow departs from the prediction systematically for many conditions in a row, an error of the model (its hysteresis compensation and its valve law at low openings) that no calibration of the reading noise on the excitation set can represent. Agner et al. report the same deterioration for their models outside the training regime.

What the sensors identify. Figure 5 and Table 4 answer the third question. With the plant's sensors 36 of 36 parameters are well identified by the quarter-of-the-prior rule and none is not, but the rule hides a gradient: the fourth section's resistance is known to 12 % of its value against 1 % or better for the other three, and the flattest direction of the information mixes it with the fourth valve's upper coefficients, the trade-off named above. Adding the four consumer differential-pressure sensors narrows that resistance 60-fold, to 0.2 %, because a pressure read across the branch alone separates the branch from the section before it. But the sensors read -0.45, -0.41, -0.30, -0.12 mWC below the model's head difference, offsets the fit absorbs as (8) allows, and even so the flows predicted with them are worse, 1.2–6.9 % on the held-out excitation conditions and 3.2–16.6 % on the realistic set: a sensor whose disagreement with the model exceeds its assumed noise adds information about the parameters and subtracts accuracy from the predictions, and its offsets are the visible part of a disagreement the model does not represent. With the flow meters alone, the third set, the first section's resistance drops out of every equation, since every consumer's path crosses it and only differences between paths remain when the pump's pressure rise is unread: it sits at its prior mean (10.00 against 10) and is the parameter classified as not identified, with 3 more only weakly. The remaining parameters are determined only relative to one another: the

Table 3: Calibration of the plant-sensor posterior's predictions, consumers 1–4: the root-mean-square error against the mean predictive standard deviation (parameter uncertainty plus the flow-meter noise of 0.044, 0.036, 0.035, 0.036 L/min calibrated on the training residuals), and how many 95 % intervals cover the measurement. On the held-out excitation conditions the two agree and the intervals cover as they should; on the realistic set the error is two to three times the interval's width.

set	n	RMSE [L/min]				mean predictive sd [L/min]				95 % intervals covering			
excitation, held-out	540	0.040	0.036	0.032	0.039	0.044	0.036	0.035	0.036	517/540	509/540	529/540	509/540
realistic	1341	0.129	0.094	0.068	0.058	0.044	0.036	0.035	0.037	546/1341	693/1341	860/1341	1033/1341

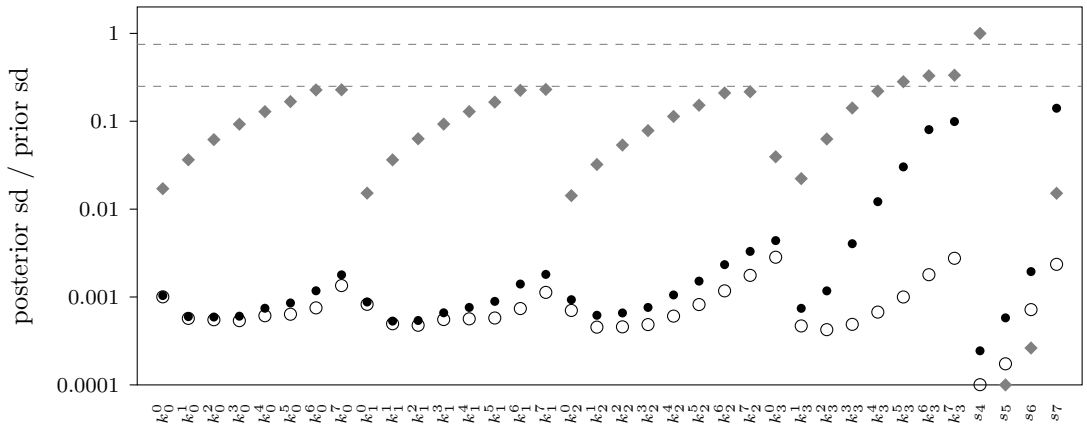


dots: measured flow; line: predicted from the pump's pressure rise and the valve positions; band: 95 % predictive interval

Figure 4: Consumer 2 on the realistic load set: the measured flow against the flow predicted from the pump's pressure rise and the commanded valve positions with the parameters inferred on the excitation set, with the 95 % predictive interval. Over the whole set the prediction's root-mean-square error is 4.0 % of the mean flow and the interval covers 52 % of the conditions: the error is systematic where the valves move slowly through positions the excitation set visited only in steps, and the interval, calibrated on the excitation set's scatter, does not contain it.

flattest direction is the common scale of every resistance and coefficient, which no flow reading can fix, and the fit settles where the priors put that scale, with valve coefficients 9.3 times the plant fit's; their small posterior standard deviations are conditional on a scale the data never saw.

Cost. 8 MAP solves, the largest over 27698 rows, produce every number in this report.



filled circles: plant sensors; open circles: plant + consumer sensors; diamonds: flow meters alone
below the lower dashed line a parameter is well identified, above the upper one it is not

Figure 5: How much each sensor set narrows each parameter: the posterior standard deviation relative to the prior's, for the 32 valve coefficients (k_i^j, valve i at knot j) and the four section resistances. With the plant's sensors 36 of 36 parameters are well identified and s_7 is the least determined; the consumer pressure sensors narrow every parameter further; the flow meters alone leave 4 weakly or not identified, s_4 at its prior.

Table 4: Identifiability of the 36 parameters under each sensor set: counts by class (posterior sd below a quarter of the prior sd is well identified, above three quarters is not), the smallest eigenvalue of the parameters' marginal information, and the direction it belongs to.

sensor set	well	weakly	not	flattest eigenvalue	flattest direction (largest components)
plant sensors	36	0	0	6.4×10^{0}	k_3^7: 0.75, k_3^6: 0.61, k_3^5: 0.23, k_3^4: 0.09
plant + consumer sensors	36	0	0	1.2×10^{4}	k_3^7: 0.80, k_3^0: 0.40, k_2^7: 0.22, k_3^6: 0.22
flow meters alone	32	3	1	2.4×10^{-1}	k_1^7: 0.31, k_0^6: 0.31, k_0^7: 0.31, k_1^6: 0.31

5 Discussion

The data are a laboratory's, and that is both the strength and the limit of the result. The line is small, its sections are identical, the valves are commanded exactly and the two experiments span the operating region deliberately; a city network offers none of these, and its calibration would face the difficulties Savic et al. (2009) list, above all too few pressure sensors for the number of pipes. The posterior here is honest on conditions like the training ones and overconfident on the realistic set, and the reason is instructive: the reading noise was calibrated on the excitation set's scatter, while the realistic set's error is systematic, a model error at low openings and slow travel that the hysteresis compensation, inherited from Agner et al., does not remove. A model-error term in the predictive variance, or a valve law with a state, would be the next step. The consumer pressure sensors show the other face of the same point: readings that disagree with the model beyond their assumed noise tighten the posterior and loosen the predictions, and their offsets should be read as a warning about the sensors or about the branch model, not as a result. The comparison with the published models is one of formulation, on the same data and the same split; the published estimator's basis of sixty ramp functions and its absolute-value objective are a different way to write a free valve curve, and the two reach similar errors where the data are rich.

6 Conclusion

This report asked three questions of a laboratory district heating line recorded by Agner et al. (2024), written as one factor graph over 1259 training load conditions. What the resistances and valve characteristics are and how well they are known: from the pump's pressure rise and the four flow meters, the first three sections' resistances and all four valve curves are inferred with standard deviations of a few percent, and the fourth section's resistance only to 12 %, its effect being nearly inseparable from the fourth valve's curve. How well the model predicts and whether its intervals are honest: held-out excitation flows are predicted to 0.9–1.7 % of the mean flow with 96 % of the 95 % intervals covering, better than the best published model on the same split (1.8–3.5 %); realistic flows to 3.1–5.6 % against 3.3–5.4 %, with only 58 % covering, because the error there is systematic and the intervals were calibrated on scatter. What each sensor set identifies: the plant's sensors everything but, weakly, one direction, the fourth section against the fourth valve; the consumer pressure sensors that direction 60-fold better, at the price of offsets and worse predictions; the flow meters alone only ratios, the first section's resistance dropping out and the scale of the rest being left to the priors. None of the constructions is new; the study leaves four exercise-sized problems in the corpus and fixes the form in which a calibration study of a heating network states its parameters, its intervals and its blind directions.

A Reproduction

```
export HYDRAULICS_DATA_DIR=...            # optional; the gate uses data/ only
python references/agner2024/fetch.py     # raw_data.zip by pinned commit
python references/agner2024/prepare.py   # the paper's prepared sets
python references/agner2024/run_reference.py   # models A, B, C refit
cd docs/terra_graph_engine/case_studies/district_heating/aalborg_parameter_posterior
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`. Without the fetched data `run.py` analyzes the committed subset under `data/` alone, which the gate test recomputes; the report is built from the full analysis. The fit has no random element.

B Data and problems

The recordings are Agner et al. (2024)'s, CC BY-SA 4.0, fetched by pinned commit and verified by hash (2001 excitation and 1341 realistic load conditions after their averaging); every twentieth condition is committed under `data/` with attribution and hashes. Under `problems/`, four directories: *valve curve from readings* (a valve's flow coefficient at two positions from pairs of flow and head readings, closed form against MAP); *flows alone fix only ratios* (two resistances in series read through flows without a pressure: the unidentified scale direction and how a single pressure reading removes it); *section and valve in series* (a pipe section feeding one valve, with and without a pressure sensor across the valve: the confounded direction and what the sensor identifies); *hysteresis compensation* (the effective position of a valve under a staircase command, and the flow error a model without the compensation makes). Each is a statement, a script that solves it with the plugin, and its results file.

References

[Savic et al. 2009] Savic, D. A., Kapelan, Z. S., Jonkergouw, P. M. R. (2009). Quo vadis water distribution model calibration? Urban Water Journal 6(1), 3–22. doi:10.1080/15730620802613380.

[Agner et al. 2024] Agner, F., Møller Jensen, C., Rantzer, A., Skovmose Kallesøe, C., Wisniewski, R. (2024). Hydraulic parameter estimation for district heating based on laboratory experiments. Energy 312, 133462. doi:10.1016/j.energy.2024.133462.

[Werner 2017] Werner, S. (2017). International review of district heating and cooling. Energy 137, 617–631. doi:10.1016/j.energy.2017.04.045.

[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.

[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.