Hydraulic parameter estimation for district heating based on laboratory experiments

Agner, Felix; Jensen, Christian Møller; Rantzer, Anders; Kallesøe, Carsten Skovmose; Wisniewski, Rafal

Published in:
Energy

DOI (link to publication from Publisher):
10.1016/j.energy.2024.133462

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

Energy 312 (2024) 133462

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Hydraulic parameter estimation for district heating based on laboratory experiments

Felix Agner [a,*], Christian Møller Jensen [b,c], Anders Rantzer [a], Carsten Skovmose Kallesøe [b,c], Rafal Wisniewski [b]

[a] Department of Automatic Control, Lund University, Ole Römers vag 1, Lund, 22363, Sweden
[b] Department of Electronic Systems, Aalborg University, Fredrik Bajers vej 7 C, Aalborg, 9220, Denmark
[c] Grundfos Holdings A/S, Poul Due Jensens vej 7, Bjerringbro, 8850, Denmark

ARTICLE INFO

Keywords:
District heating
Experiment
Hydraulic
Model
Parameter estimation

ABSTRACT

In this paper we consider calibration of hydraulic models for district heating networks based on operational data. We extend previous theoretical work on the topic to handle real-world complications, namely unknown valve characteristics and hysteresis. We generate two datasets in the Smart Water Infrastructure Laboratory in Aalborg, Denmark, on which we evaluate the proposed procedure. In the first data set the system is controlled in such a way to excite all operational modes in terms of combinations of valve set-points. Here the best performing model predicted volume flow rates within roughly 5 and 10 % deviation from the mean volume flow rate for the consumer with the highest and lowest mean volume flow rates respectively. This performance was met in the majority of the operational region. In the second data set, the system was controlled in order to mimic real load curves. The model trained on this data set performed similarly well when evaluated on data in the operational range represented in the training data. However, the model performance deteriorated when evaluated on data which was not represented in the training data.

1. Introduction

The energy systems of the near-future face many challenges such as higher penetration of renewable energy production, and large and fluctuating energy demand. To face these challenges, a promising outlook is to introduce a stronger coupling between smart grids, smart thermal networks and smart gas networks. In this proposed architecture, the emerging 4th generation of district heating networks, characterized by decentralized heat production, significant integration with smart power and gas networks, and low distribution temperatures, is likely to take a central role [1]. Operating these new systems come with many challenges related to control and operation [2]. To face these new challenges, smart systems and techniques for modeling and simulating the systems will be key. While modeling and simulating the thermodynamics of district heating network for operational optimization now has a now rather long tradition [3], recent works have also started to delve deeper into the hydraulics of district heating networks. Hydraulic models are useful for simulating the hydraulic state of district heating networks [4–6]. This information allows system operators to optimally control pumps at maximum energy efficiency [7–10] or distribute heat in a fair fashion when the heat demand exceeds the available network capacity [11]. In essence, hydraulic models describe the relationship between pressures and volume flow rates for hydraulic system components such as valves and pipes. Traditionally, such models have consisted of fully white-box models. For pipes this implies models based on friction coefficients and pipe dimensions. For valves, it has implied manufacturer-provided characteristic curves. While the emerging energy systems pose many tough engineering problems, they also open the door to improving these models by tuning them based on operational data. This is due to the expected increase of smart metering in modern systems [2]. This data-driven approach holds many advantages compared to traditional methods. Data driven models can be updated to better fit the observed behavior of a component under operation and they can be adapted over time to compensate for degradation. On the contrary, gathering and maintaining correct information to build traditional white-box models for hydraulic components, e.g. pipe dimensions and friction coefficients, can be tedious and error-prone. While one could consider fully black-box models for specific purposes, there are advantages to maintaining a component-based grey-box model structure. This approach maintains physical interpretability and allows the model to be used in applications other than simulation, such as fault detection [12].

* Corresponding author.
E-mail address: felix.agner@control.lth.se (F. Agner).

https://doi.org/10.1016/j.energy.2024.133462
Received 7 July 2024; Received in revised form 9 October 2024; Accepted 12 October 2024
Available online 21 October 2024

The notion of data-driven model calibration has been studied longer in the field of water distribution networks, as reviewed in [13], than in the field of district heating. However, district heating networks typically have a distinct structure from water distribution networks, which can be exploited in modeling efforts: district heating networks form closed hydraulic circuits, whereas water distribution networks end in open outlets at the points of consumption. Several works aimed particularly at modeling in the district heating domain have been published in recent years. The authors of [9] performed pipe resistance estimation and subsequently used the identified model to demonstrate effectiveness of an advanced control strategy in simulation. They estimated the resistance of all pipes in the supply-and-return-lines separately, based on measurement of pressure at the supply-and-return connections of each substation. In [14], the authors investigated the same strategy while demonstrating a way to account for meshed structures in the system by successively closing sets of loop-generating pipes. In [15], yet a third simulation study with this approach was performed, where the authors now focused on reducing the number of required measurements in the system. In [16], an extension of the framework was proposed where estimation of valve characteristics was included through exploiting the closed hydraulic circuit structure of district heating networks. A common trend among these works is that they are based on mathematical analysis of the networks along with simulations. Zheng et al. on the other hand perform pipe parameter identification and compare their model to real operational data from a real system in Tianjin, China [17]. This is the first and only work to the authors' knowledge that performs a study on real data. However, this study is restricted to models of pipe parameters, excluding valve models. They are also unable to share their data with the scientific community. Studying data from real systems is highly important, as many issues which are not visible in simulation can emerge. Measurement noise causes an obvious issue, but also problems based on unmodeled properties of the system. A clear example of such an aspect which is not covered in previous literature on hydraulic district heating model calibration, but which is a well-known issue to engineers working on physical systems, is valve hysteresis [18].

In this work, we perform experiments in a laboratory setup as a step towards bridging the gap between the theoretical literature and real application. The study was performed in the Smart Water Infrastructure Laboratory (SWIL) at Aalborg University [19]. We present the following contributions.

1. We build a test-bed simulating the hydraulic properties of four buildings in a line-structured district heating network.
2. We generate two data-sets. Firstly a data-set designed to excite all operational modes of the system. Secondly a data-set based on real district heating load curves. These data sets are open and available for future work.
3. We extend the modeling approach presented in [16] to handle real world complications. Namely, valve characteristics are not known a priori, and valves are subject to hysteresis.

The paper is organized in the following way. We formally introduce the problem and the considered parameter estimation framework in Section 2. We then go into detail regarding the specific model structures we consider in Section 3. The experimental setup is described in Section 4. We present and discuss the results of our study in Section 5. We conclude the paper in Section 6 and present topics for future work in Section 7.

1.1. Notation

We denote volume flow rates (l/min) by the variable q, pressures (mH2O) by the variable p and valve set-points by the variable v. We use subscripts to connect the measured values to a specific component, e.g. q_i is the volume flow rate measured in the component of index i. For a matrix M we denote $M_{i,j}$ to be the entry of M at row i and column j. We denote the ramp function as

$$\mathrm{ramp}_a^b(x) = \begin{cases} 0, & \text{if } x \le a \\ (x-a)/(b-a), & \text{if } a < x \le b \\ 1, & \text{else} \end{cases} \tag{1}$$

where we assume that $b > a$. We denote a column vector of all 1's of size n as $\mathbf{1}_n$ and an n by m matrix of all zeros as $\mathbf{0}_{n\times m}$.

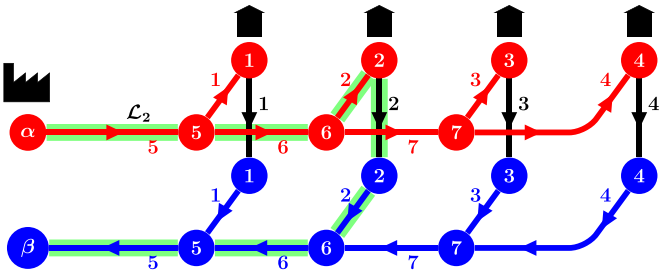


Fig. 1. A schematic example of a tree-structured district heating network. The nodes $\mathcal{N} = \{\alpha, 1, \dots, 7\}$ are connected by edges $\mathcal{E} = \{1, \dots, 7\}$. The supply-and-return-networks are symmetrical except for their edges having opposite direction. Valves $\mathcal{V} = \{1, \dots, 4\}$ connect the supply and return lines. Root nodes α and β in the supply-and-return networks respectively connect the branch either to a production facility, or a greater part of the network. The path $\mathcal{L}_2 = \{2, 5, 6\}$ highlighted in green contains the edges leading from α to β via valve 2.

2. Hydraulic parameter estimation

We consider the problem of finding parameters for hydraulic models of district heating networks, based on operational data corresponding to load conditions measured at times $t = 1, \dots, T$. We restrict ourselves to networks that have a tree-shaped structure, an example of which can be seen in Fig. 1. We model the supply-side network as a structured graph $\mathcal{G} = (\mathcal{E}, \mathcal{N})$ which we assume to have a tree-structure. $\mathcal{E}$ is the set of edges which represent pipes, and $\mathcal{N}$ is the set of nodes where these pipes connect. This graph-based perspective is a standard way of representing district heating hydraulics [20]. The supply-and-return networks are assumed to be symmetrical and therefore the return network can be represented by the same graph $\mathcal{G}$, with two exceptions. Firstly, we assume that the edges have opposite direction in the supply and return networks. Hence as the water flows out through the supply network and returns via the return network, the sign of the flow rates through these edges will be the same. Secondly, we consider one node in the graph to be the root of the tree, which we denote α in the supply network and β in the return network. This root can represent e.g. a production facility, or a connection to a larger network to which the considered network is a small sub-network. A subset $\mathcal{V} \subset \mathcal{N}$ of the nodes correspond to consumers. They connect the supply and return lines via control valves. In the considered example, $V = \{1, 2, 3, 4\}$. In a real setting, a single consumer may in fact be represented by two (or more) control valves, as domestic hot water and space heating are typically hydraulically separated [21, p. 365]. For each valve $i \in \mathcal{V}$, there is a set $\mathcal{L}_i$ of edges corresponding to a path from α to β through valve i. For instance, $\mathcal{L}_2 = \{2, 5, 6\}$ in our example as highlighted in Fig. 1.

For the considered parameter estimation protocol, we assume that we can measure the root pressure difference $\Delta p_0 = p_\alpha - p_\beta$. We additionally assume that we can log the control valve set-points v_i for each consumer and the volume flow rates q_i through each valve $i \in \mathcal{V}$.

The networks are also assumed to have the following property.

Assumption 1. For all $i \in \mathcal{N}$, either $i \in \mathcal{V}$ or i is connected to at least 3 edges.

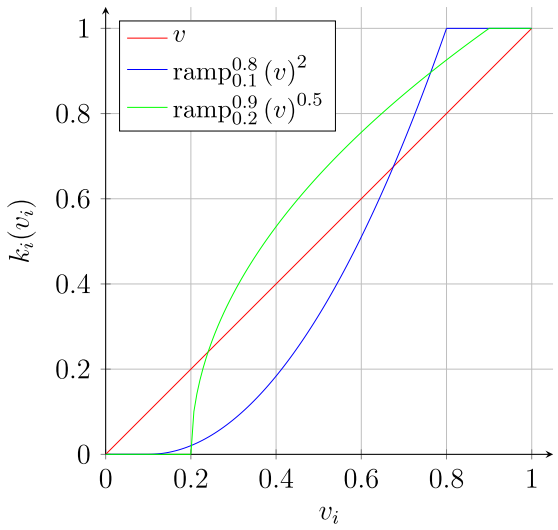


Fig. 2. Examples of valve characteristics $k_i(v_i)$.

In practice, this assumption means that no pipes in the model are placed in direct series. Any physical interconnection of pipes in direct series would be represented by one single, equivalent pipe in the model. For instance, it would be reasonable to expect that a consumer substation has a physical interconnection of several pipe sections in series with the control valve and a heat exchanger. The heat exchanger and the pipes would then be concatenated into one single pipe in the model.

With the considered model in mind, we have two types of hydraulic components; *pipes* and *valves*. For each such component, we aim to parameterize a model which maps the volume flow rate through the component to the pressure difference.

2.1. Pipe parameterization

For pipes, we consider models stemming from the Darcy–Weisbach equation. If the pipe corresponds to edge i, which leads from node j to node k, the relation between the pressure loss $\Delta p_i = p_j - p_k$ and the volume flow rate q_i along the direction of the pipe is given by

$$\Delta p_i = s_i q_i |q_i|^{\gamma}. \tag{2}$$

Here s_i is the model parameter which our framework aims to estimate from data, typically referred to as the pipes hydraulic resistance. γ is an exponent which in the literature typically takes the value 1 or slightly lower e.g. 0.87. We limit the scope of this study to consider the assumption $\gamma = 1$. We make the following assumption regarding the parameters s_i.

Assumption 2. The pipe resistance s_i is equal for the corresponding supply-and-return-lines.

It was shown in [16] that this assumption along with Assumption 1 is sufficient for unique identification of model parameters. In practice, it is unlikely in that this assumption holds exactly due to slight differences in the pipes. However, we cannot expect to find unique parameters for these pipes with our methodology, given that we only measure pressure at two points in the network (α and β). However, for many model applications such as estimating volume flow rates (see Appendix A), only the sum of the parameters for the supply-and-return-lines is necessary.

2.2. Valve parameterization

The second and final hydraulic model component type we consider is valves. For valve i which connects node i in the supply network to the return network, Δp_i is the pressure difference between these nodes. Δp_i is given by the volume flow rate q_i and the valve set-point v_i which varies from fully closed ($v_i = 0$) to fully open ($v_i = 1$) in the following way.

$$q_i = K_{v,i} k_i(v_i)\sqrt{\Delta p_i} \implies \Delta p_i = \frac{1}{K_{v,i}^2 k_i(v_i)^2} q_i^2 \tag{3}$$

where $K_{v,i}$ represents the hydraulic admittance of the valve when it is fully open. This is often referred to as the flow coefficient of the valve, or the k_{vs} or c_{vs} value in Nordic or English literature respectively [21, p. 399]. k_i represents the valve characteristics, i.e. the mapping between the valve set-point and the valve admittance. The valve characteristic function may be linear, e.g. $k_i(v_i) = v_i$. It may however also have other shapes. One common valve characteristic is quick opening, where at low set-points, a small increase in valve set-point will lead to a large increase in flow-through. Another common characteristic is equal percentage which is in a sense the opposite of quick opening. Here it is expected that the valve needs to provide a larger change in volume flow rate in response to small changes in set-point at higher levels of opening [18, pp. 82–83]. Additionally, one might consider valve characteristics where the valve fully closes at a value $v > 0$ and fully opens at a value $v < 1$, to ensure that the full capacity of the valve is used at high set-points, and that the valve is fully closed at near-0 set-points. Fig. 2 shows three examples of $k_i(\cdot)$, corresponding to a linear valve ($k_i(v_i) = v_i$), an equal percentage valve which fully closes at 0.1 and fully opens at 0.8 ($k_i(v_i) = \text{ramp}_{0.1}^{0.8}(v_i)^2$), and a quick-opening valve which fully closes at 0.2 and fully opens at 0.9 ($k_i(v_i) = \text{ramp}_{0.2}^{0.9}(v_i)^{0.5}$). As we cannot necessarily know the valve characteristics of all of the valves we aim to parameterize, we propose a linear combination of typical characteristics, aimed to capture (3):

$$\Delta p_i = \left(\frac{\theta_{i,1}}{k_1(v_i)^2} + \frac{\theta_{i,2}}{k_2(v_i)^2} + \cdots + \frac{\theta_{i,K}}{k_K(v_i)^2} \right) q_i^2. \tag{4}$$

Here $k_k(\cdot)$, $k = 1, \ldots, K$ correspond to different possible valve characteristics and $\theta_{i,k}$ corresponds to a tunable parameter for valve i in relation to the valve characteristic $k_k(\cdot)$. Somehow $1/k_k(\cdot)^2$) are basis of all possible characteristics of a valve in a K-dimensional space. Hence while the actual valve characteristic may not be perfectly captured by any of the functions $k_k(\cdot)$, a linear combination of them may be able to replicate the actual behavior of the valve.

2.3. Parameter estimation

To estimate the values of s_i (for pipes) and $\theta_{i,k}$ (for valves), we assume that we have access to data from T different steady state load conditions. We first consider one such load condition at time t. Here we measure the pressure difference at the tree root $\Delta p_0(t) = p_\alpha(t) - p_\beta(t)$, the valve flow rates $q_i(t)$, and valve set-points $v_i(t)$ where $i \in \mathcal{V}$.

We can use the valve volume flow rates $q_i(t)$ to calculate the volume flow rates in all of the pipes. To do so, we first define the basic incidence matrix $B \in \mathbb{R}^{(n_{\mathcal{V}}-1)\times n_{\mathcal{E}}}$ for our considered graph representation of the network $\mathcal{G}$. Here $n_{\mathcal{E}}$ and $n_{\mathcal{V}}$ are the number of edges and nodes respectively. The elements of B are given by

$$B_{i,j} = \begin{cases} 1, & \text{if edge } j \text{ leads to node } i, \\ -1, & \text{if edge } j \text{ leads from node } i, \\ 0, & \text{else.} \end{cases} \tag{5}$$

We omit the row corresponding to the root node α or β, which together with the tree-structure of $\mathcal{G}$ ensures the invertibility of B. Hence we can find the vector $q_{\mathcal{E}}$ corresponding to the vector of all flow rates in $\mathcal{E}$ as

$$q_{\mathcal{E}}(t) = B^{-1} q_{\mathcal{N}}(t) \tag{6}$$

where $q_{\mathcal{N}}(t) \in \mathbb{R}^{n_{\mathcal{N}}-1}$ corresponds to the total volume flow rate out of each node $i \in \mathcal{N} \setminus \alpha$. I.e.

$$q_{\mathcal{N},i}(t) = \begin{cases} q_i(t), & \text{if } i \in \mathcal{V}, \\ 0, & \text{else.} \end{cases} \tag{7}$$

With this information, the volume flow rates through all pipes and valves are known. We can now consider one $i \in \mathcal{V}$, and the corresponding path $\mathcal{L}_i$ from α to β through valve i. It must hold that the total pressure difference $\Delta p_0(t)$ between α and β corresponds to the pressure drop along each traversed component in $\mathcal{L}_i$, i.e.

$$\begin{aligned}\Delta p_0(t) &= \Delta p_i(t) + 2\sum_{j\in\mathcal{L}_i} \Delta p_j(t) \\ &= \left(\frac{\theta_{i,1}}{k_1(v_i(t))^2} + \frac{\theta_{i,2}}{k_2(v_i(t))^2} + \cdots + \frac{\theta_{i,K}}{k_K(v_i(t))^2}\right) q_i(t)^2 + 2\sum_{j\in\mathcal{L}_i} s_j q_j(t)|q_j(t)|.\end{aligned} \tag{8}$$

Here $\Delta p_j(t)$ is equal for the edges in the supply-and-return networks due to Assumption 2 and symmetry, yielding the factor 2 in front of the sum. We can use all paths $\mathcal{L}_i$ where $i = 1, \dots, n_\mathcal{V}$ to formulate a system of equations

$$\mathbf{1}_{n_\mathcal{V}} \Delta p_0(t) = \begin{bmatrix} F_1(t) & \mathbf{0}_{1\times K} & \dots & \mathbf{0}_{1\times K} \\ \mathbf{0}_{1\times K} & F_2(t) & \dots & \mathbf{0}_{1\times K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1\times K} & \mathbf{0}_{1\times K} & \dots & F_{n_\mathcal{V}}(t) \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{n_\mathcal{V}} \end{bmatrix} + G(t)s = F(t)\theta + G(t)s. \tag{9}$$

Here $G(t) \in \mathbb{R}^{n_\mathcal{L}\times n_\mathcal{E}}$ is a matrix defined as

$$G_{i,j} = \begin{cases} 2q_j(t)|q_j(t)| & \text{if edge } j \in \mathcal{L}_i, \\ 0 & \text{else.} \end{cases} \tag{10}$$

The vectors θ_i for $i = 1, \dots, n_\mathcal{V}$ gather the parameters $\theta_{i,k}$ for $k = 1, \dots, K$, and $F_i(t) \in \mathbb{R}^{1\times K}$ is given by

$$F_i(t) = \begin{bmatrix} \frac{q_i(t)^2}{k_1(v_i(t))^2} & \frac{q_i(t)^2}{k_2(v_i(t))^2} & \dots & \frac{q_i(t)^2}{k_K(v_i(t))^2} \end{bmatrix}. \tag{11}$$

We now move from considering only one load condition at time t to concatenating the data from all T load conditions into one large system of equations

$$\Phi \begin{bmatrix} \theta \\ s \end{bmatrix} = y \tag{12}$$

where

$$\Phi = \begin{bmatrix} F(1) & G(1) \\ F(2) & G(2) \\ \vdots & \vdots \\ F(T) & G(T) \end{bmatrix}, \quad y = \begin{bmatrix} \mathbf{1}_{n_\mathcal{V}} \Delta p_0(1) \\ \mathbf{1}_{n_\mathcal{V}} \Delta p_0(2) \\ \vdots \\ \mathbf{1}_{n_\mathcal{V}} \Delta p_0(T) \end{bmatrix}. \tag{13}$$

This way Φ is a data-matrix constructed by measuring flow rates and valve set-points and y consists of measured root pressure measurements. Once Φ and y are constructed, there are many ways to choose s and θ to fit the observed data. We choose to fit the parameters numerically as the solution to

$$\underset{s,\theta}{\text{minimize}} \quad \left\| \Phi \begin{bmatrix} s \\ \theta \end{bmatrix} - y \right\|_1 \tag{14a}$$

$$\text{subject to } s \geq 0 \tag{14b}$$

$$\theta \geq 0. \tag{14c}$$

The choice of using $\|\cdot\|_1$ in (14a) is based on the general robustness to outliers typically provided by this cost function. The positivity constraints Eq. (14b)–(14c) enforce physical feasibility of the model. We solved (14a) using the python toolbox `CVXPY` [22,23] using the `SciPy` solver. A natural extension of (14a) is to introduce parameter regularization to (14a). This would reduce the risk of overfitting and promote sparseness in the model depending on the choice of regularization employed. Our investigation has however showed that the models we found are already sparse, and therefore leave an investigation of the impact of regularization outside the scope of this paper.

3. Model structures and data preprocessing

In this section we detail three different model structures which we subsequently compare on our experimental data:

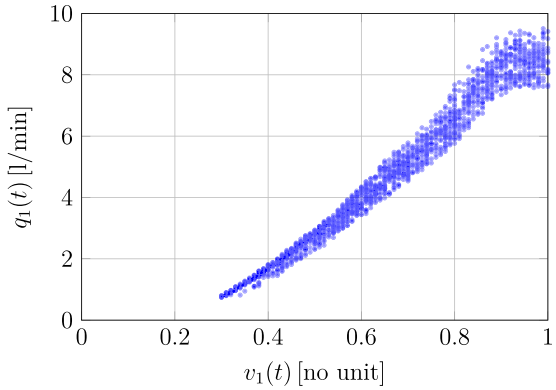


Fig. 3. Scatter-plot of flow rates $q_1(t)$ against valve set-points $v_1(t)$ for valve 1 in the experiment setup to be described in the subsequent section.

A: A simple model with a linear valve curve assumption.
B: Parameterized, nonlinear valve curves.
C: The same valve curves as B, but including data preprocessing for hysteresis compensation.

3.1. Model A: Naive valve curve parameterization

Our first and most simple model equates to the model considered in [16]. Here we assume that all valves are fully linear and thus can be parameterized by only one function $k(\cdot)$:

$$k(v_i) = v_i. \tag{15}$$

3.2. Model B: Enhanced valve curve parameterization

To account for nonlinear valve characteristics, we consider also a larger number of parameterization functions $k_k(\cdot)$. To make an educated guess about reasonable valve curve shapes, Fig. 3 shows a scatter plot of $q_1(t)$ over $v_1(t)$ for valve 1 in the experimental setup to be described later. While this scatter-plot does not account for changes in differential pressure over the valve, we can clearly see that a linear curve from 0 to 1 does not capture this shape. Instead, it appears that the curve should have a sub-linear curve corresponding more to equal percentage characteristics. Additionally, the valve is clearly fully open already around $v_1 \approx 0.9$, and fully closed around $v_1 \approx 0.2$ To handle these nonlinear patterns, we consider valve characteristic functions on the form

$$k_k(v_i) = \text{ramp}_{a_k}^{b_k}\left(v_i\right)^{c_k}. \tag{16}$$

for different choices of a_k, b_k and c_k. Based on our intuition from Fig. 3, we considered all such combinations corresponding to

$$a_k \in [0.10, 0.15, 0.20, 0.25], \tag{17}$$

$$b_k \in [0.80, 0.85, 0.90, 0.95, 1.0], \tag{18}$$

$$c_k \in [1.0, 1.25, 1.5], \tag{19}$$

resulting in $K = 60$ different valve parameterization functions $k_k(\cdot)$ for each valve. We chose the values for a_k and b_k as heuristically reasonable values for full opening and closing, and the values of c_k as values corresponding to reasonable curves for equal percentage valves, which was deemed appropriate based on Fig. 3.

3.3. Model C: Hysteresis compensation

Our third and most sophisticated model uses the same valve characterization as model B. Model C however accounts for the effect of hysteresis. This is an effect which typically manifests in control valves,

causing an issue where the valve spindle does not move for small changes in the set-point given to the valve [18, p. 88]. We account for this issue in the following way. Instead of using the valve set-points $v_i(t)$ for model fitting and validation, we utilize filtered versions $\hat{v}_i(t)$ which are meant to estimate the actual valve spindle position. We first define a hand-tuned parameter $\delta = 0.015$. The value 0.015 was found heuristically by testing the values $0.05, 0.10, 0.15, 0.20$ and then choosing the best one. The spindle is assumed to remain still whenever the set-point moves a distance less than δ from the current estimated spindle position. Once the reference moves outside of this range, the estimated spindle position will lag behind by a distance δ. These assumptions are captured in the following filter:

$$\hat{v}_i(t) = \begin{cases} \hat{v}_i(t-1) & \text{if } \left|v_i(t) - \hat{v}_i(t-1)\right| \leq \delta \\ v_i(t) - \delta & \text{if } v_i(t) \geq \hat{v}_i(t-1) + \delta \\ v_i(t) + \delta & \text{else.} \end{cases} \tag{20}$$

The filtered time series are initiated with the first measured valve set-point $\hat{v}_i(0) = v_i(0)$.

4. Experiment description

Our experiment setup represents four domestic buildings connected through a line-structured district heating network as seen in Fig. 1. This is the same structure that we used as an example in Section 2. The four consumers are connected to a single production facility. One could also imagine that the root node in the laboratory setup represents a connection between the sub-branch of these four consumer to a larger network. We will now describe the laboratory equipment used to represent the production facility, the pumping, and the consumers.

4.1. Equipment

The experiments were conducted in Aalborg University's Smart Water Infrastructure Lab (SWIL). SWIL is a state-of-the-art water and district heating laboratory that consists of configurable modules which can be used to build a wide variety of experimental setups [19]. The available module types are:

- Pumping stations
- Consumer stations
- Piping stations
- Heating stations

Examples of consumer and piping stations are seen in Fig. 5. Pumping and consumer stations are particularly configurable and may be used to emulate e.g. gravity sewers or elevated reservoirs. The experimental network corresponding to the schematic Fig. 1 consists of a pumping station, two piping stations, and two consumer stations as seen on Fig. 4. The consumer stations represent two consumers (buildings) each. This corresponds to internal piping, a control valve, and a heat exchanger in series connection.

4.2. Measurements

Flows and pressure were measured using the piping station sensors, respectively an Endress & Hauser Proline Promag 10 and Grundfos Direct Sensor RPI+T 0 – 1.6, connected to Beckhoff I/O modules. The data was collected at 1 Hz and transmitted from the laboratory modules to a central control unit (CCU). The laboratory modules are controlled by soft PLCs comprising CodeSys software running on a Raspberry Pi. The CCU communicates with the modules over Modbus TCP through a Simulink simulation using the MATLAB Industrial Communications Toolbox and the Simulink Real-Time Pacer [24], and raw measurements are converted to respectively l/min and mH2O at this stage. The consumer volume flow rates $q_1, \ldots, q_4$ were subsequently calculated as

$$q_1 = q_{\mathrm{FT\_1}} - q_{\mathrm{FT\_2}}, \tag{21}$$

$$q_2 = q_{\mathrm{FT\_2}} - q_{\mathrm{FT\_3}}, \tag{22}$$

$$q_3 = q_{\mathrm{FT\_3}} - q_{\mathrm{FT\_4}}, \tag{23}$$

$$q_4 = q_{\mathrm{FT\_4}} \tag{24}$$

where e.g. $q_{\mathrm{FT\_1}}$ is the volume flow rate measured in sensor FT_1. The pressure measurements PT_1 and PT_2 were used to represent the network root pressures p_α and p_β respectively.

4.3. Data sets

We generated two experimental datasets, henceforth referred to as the *exciting* and *realistic* data sets.

In the *exciting* data set, we kept the pump working at full capacity and set the valve set-points randomly to individually and uniformly sampled values in the interval $[0.3, 1.0]$ every 40 s. The lower bound of 0.3 corresponds to valves which are in practice almost fully closed. This bound was included to avoid that a large part of the data corresponds to closed valves. During the full experiment, the pump was operating constantly at full capacity. In post-processing, the first 10 s in each 40 s interval was removed in order to ignore the transient phase of each step. The mean of remaining 30 s was used to generate a data point for later model fitting and validation. This process is shown in Fig. 6.

The exciting data set was produced to give us the best possible chance of accurately estimating the system parameters, but would be hard to replicate in a real district heating network. In the *realistic* data set, we instead designed the system to emulate the recorded volume flow rates in four residential buildings in Nuremberg, Germany, from 2022. We used two weeks of hourly mean values. We scaled the magnitude of the recorded flow rates such that they would fit in the operational range of our experimental setup. In the experiment, this data served as the reference volume flow rate for each consumer, which we tracked using the control valves of the consumer units. Each control valve was equipped with hand-tuned PID controller for this purpose. The pump was operating constantly at full capacity during the experiment to ensure sufficient system pressurization. For the purpose of experiment tractability, we scaled down the time such that 1 h of reference data corresponded to 2 min in the experiment. The resulting *exciting* volume flow rates are shown in Fig. 7. We can see that the consumption of these consumers follows a rather typical pattern. There are slow dynamics, governed by the demand for space heating driven by the outdoor temperature, along with faster intra-day peaks driven by domestic hot water consumption. We used mean values of 30 s of experimental data for the subsequent model fitting and validation which corresponds to 15 min mean value measurements in the real district heating setting. The reference volume flow rates correspond to the combined load for domestic hot water and space heating. In a real system, these loads are often separated into two different hydraulic loops controlled by separate valves [21, p. 365], but such a separation would be infeasible in our small laboratory setup.

4.4. Model evaluation

We choose to evaluate the predictive power of the model from a given set of control inputs (valve set-points v_i and differential pressure Δp_0 provided by the pump) to system outputs (flow rates though the four control valves representing consumers). Appendix A describes how to use the parameterized models to calculate estimated volume flow rates $\hat{q}_i$ in this way. We then investigate the prediction errors $e_i = q_i - \hat{q}_i$ to judge the model performance. This is a metric which is relevant when using the model to design flow rate controllers. We can note that this is not the objective for which the parameters are tuned. The training objective (14a) rather reflects the ability of the model to predict pressure drops given the valve set-points and flow rates in the system. The component-wise modeling of the network is what allows this flexibility in model application. This multi-purpose

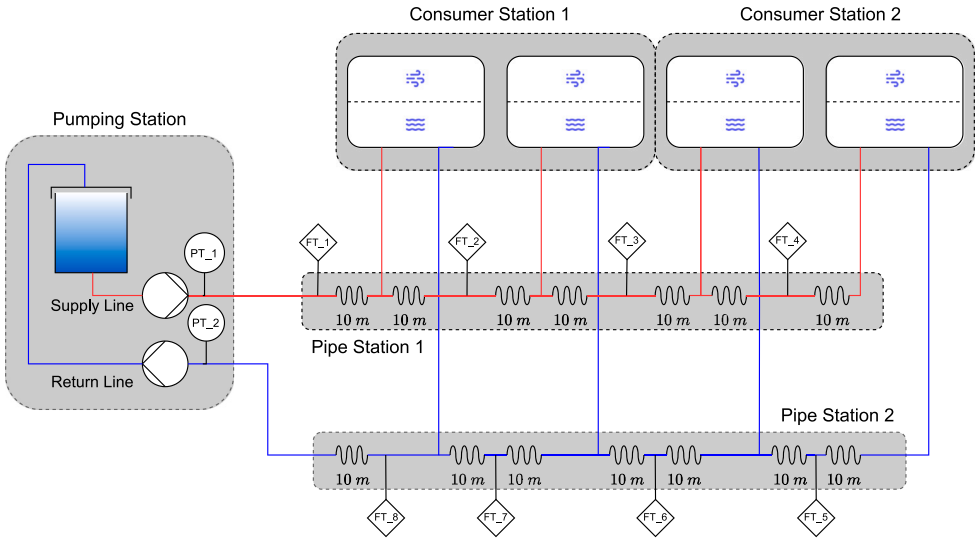


Fig. 4. Diagram of the experimental setup used to recreate Fig. 1. PT and FT denote respectively flow and pressure transmitters, while the two blocks in each consumer station are water-to-air heat exchangers. Each heat exchanger is connected in series with a control valve used to regulate the volume flow rate.

(a) On the left are two pipe stations. These stations measure volume flow rates and pressures, and can be configured for pipe sections of varying lengths. On the right is a pumping station, where water is pumped from a water reservoir into the system, until finally returning to the reservoir.

(b) Two consumer stations, which in total simulate four consumers. The bottom sections are connected to the pipe units. Water then flows up and through the water-to-air heat exchangers at the top, via a control valve.

Fig. 5. Images from the laboratory setup.

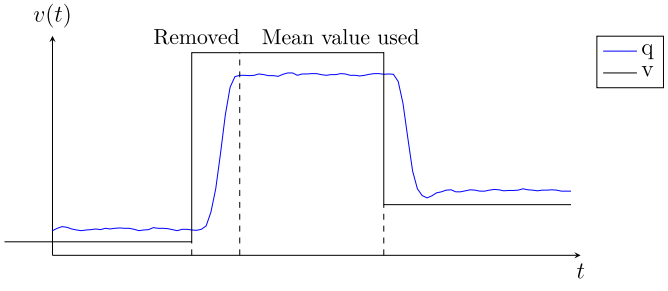


Fig. 6. Illustration of the method used to generate the *exciting* data set. Valve set-points are randomly generated every 40 s. The subsequent 10 s of data is discarded to remove transient effects, and the mean of the remaining 30 s of data is logged as a data point. Valve set-points are lower-bounded by 0.3.

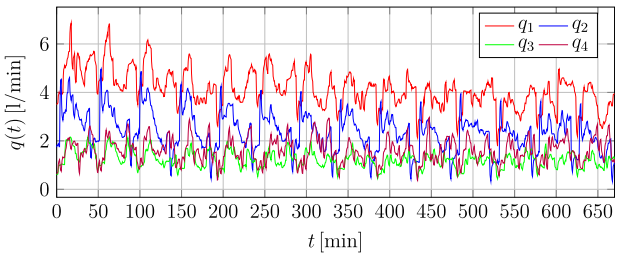


Fig. 7. Flow rates $q_i(t)$ in the realistic data set, recorded during the experiment. 48 min corresponds to 24 h of real world time, in which we can see typical daily consumption patterns.

capacity is testimony to the value of considering grey-box component-models of this type over black-box models such as neural networks or regression models without structure considerations. There are other relevant choices of evaluation that could also be considered. In [16], ground-truth parameter value comparison was used. This is however clearly infeasible in the real setting where no ground truth is available.

5. Results

The following section details the results of the experiments. Firstly the model structures A, B and C trained on the exciting data set are investigated. Afterwards we evaluate the best performing model structure C trained on the realistic data set.

5.1. Exciting training data

Table 1 shows the resulting pipe parameters s_i using the different models. Pipes 1–4 lead to only one of the four consumers, whereas pipes 5–7 are shared between the consumers in the sense that the e.g. the volume flow rate for all four consumers comes through pipe 5. Of interest is that the resistances s_1, s_2 and s_3 were estimated as 0 for all three models trained on the exciting data set. One explanation for this is that our data does not represent the section of the valve curve where the valve is fully closed. To see why this matters, consider Fig. 8(a). Here we see a hydraulic schematic representing one consumer in the network, i.e. two symmetric pipes (here represented by a total resistance s) and a valve (with curve $K_v k(v)$). The relationship between $\Delta p = p_i - p_j$, i.e. the pressure difference between supply-and-return-lines where the consumer is connected, and the flow rate q for the consumer will be

$$\Delta p = \left(s + \frac{1}{K_v^2 k(v)^2} \right) q^2 \tag{25}$$

and therefore

$$q = \frac{K_v k(v)}{\sqrt{s K_v^2 k(v)^2 + 1}} \sqrt{\Delta p}. \tag{26}$$

We can denote $\hat{k}(v) = \frac{K_v k(v)}{\sqrt{s K_v^2 k(v)^2 + 1}}$. Consider two such choices of $\hat{k}(v)$: Firstly, $\hat{k}_1(v)$ where $s = 1.5$, $k(v) = \text{ramp}_{0.1}^{0.9}(v)^{1.5}$ and $K_v = 1$. Secondly, $\hat{k}_2(v)$ with $s = 0$ (no pipe), $k(v) = \text{ramp}_{0.22}^{0.9}(v)^{0.75}$, $K_v = 0.63$. Fig. 8(b) shows the plots of $\hat{k}_1(v)$ and $\hat{k}_2(v)$. Here we see that if we consider only valve-positions above $v = 0.3$, the curves are almost equivalent even though one of them has no pipe resistance. In the same way, due to the fact that we chose to omit valve openings below 0.3 in the training data, the model cannot clearly discriminate between the valve curve parameters and the pipe parameters. This is an important realization, demonstrating that effective data sets should include also valve set-points in the lower range of operation. There are a couple of possible explanations as to why $s_4 \neq 0$ for models B and C while the other three are 0. Firstly pipes 1–3 should be equivalent due to the experiment configuration, with pipe 4 being longer than the others and thus more resistive. Secondly, since consumer 4 is the furthest from the pump, it also experiences the lowest differential pressure, which causes the volume flow rate to occasionally drop to 0. This could influence the estimation procedure towards $s_4 \neq 0$. Pipes 5, 6 and 7 do not experience the same issue as discussed with regards to Fig. 8, most likely because they are not in strict series with a valve. Pipes 6 and 7 should be similar in size. However, in models B and C, pipe 7 appears roughly a factor 5 longer than pipe 6, indicating that either the model does not capture the contribution of the pipes fully, or there are other reasons such as small twists and turns in the piping causing a difference between pipes 5 and 6.

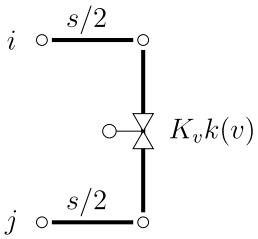
The resulting valve curve models become very sparse, an effect likely caused by the positivity constraint (14c) which prevents over-fitting. Even though there are in total $K = 60$ parameters $\theta_{i,k}$ for each valve i, resulting in 240 total valve curve parameters, only one of the valve curves has more than 5 non-zero parameters $\theta_{i,k}$. In total, model B has 18 parameters $\theta_{i,k} > 0$ and model C has 16. This model sparsity motivates the decision to omit parameter regularization from the scope of this paper. The full equation for each valve curve and each model as per (3) is provided in Appendix B. We can also investigate the valve characteristics visually, which yields Fig. 9. As all four consumers are constructed with equal equipment, we would expect the four valve characteristic curves to be equivalent. However, this is not the case. Rather, valves 1 and 2 have similar characteristics, whereas valves 3 and 4 have much lower curves. This means that given equal differential pressure and valve set-points, the model predicts that valve 4 would yield a lower volume flow rate than the other valves. One possible explanation for this behavior is that the exponent γ in (2) and q_i^2 in (3) do not perfectly capture the real behavior of the system. If these exponents are too high, it means that the expected pressure drops at high flow rates will be over-estimated. Clearly, the volume flow rates and valve openings are strongly correlated (see Fig. 3), hence the model can counter-act this effect by increasing the admittance at high valve set-point values. The reason this could make valve 4 stand out from the others is that the average volume flow rates over valve 4 are much lower than the other valves (the mean volume flow rates are $\bar{q}_1 = 4.60$ l/min, $\bar{q}_2 = 4.34$ l/min, $\bar{q}_3 = 2.91$ l/min, and $\bar{q}_4 = 2.11$ l/min respectively) and thus the model does not need to over-estimate the admittance at high flow rates to the same extent.

Fig. 10 shows the prediction errors $e_i(t) = q_i(t) - \hat{q}_i(t)$ plotted over the corresponding valve set-points $v_i(t)$. Each column represents one model structure A, B or C, and each row represents one valve index $i = 1, \dots, 4$. Fig. 10 shows that A has a strong correlation between prediction errors and valve set-points, which means that this naive parameterization of the valve curves is clearly insufficient. Both models B and C have a much lower correlation between valve set-points and prediction errors, with a few exceptions. Valves 1–3 display strong correlation between e_i and v_i around $v_i \approx 0.85$ which is the inflection point where the valve enters the fully closed position. It is possible that further tuning of the chosen parameterization $k_k(\cdot)$ could remove this issue. Valve 4 shows a cluster of outlier errors around $v_4 \approx 0.3$. We can find an explanation in Fig. 11 where we see the flow rate $q_4(t)$ scattered over the valve set-point $v_4(t)$ in the exciting data set. We can note that there are barely any recorded volume flow rates between 0 and 0.3 l/min. This means that when the valve is almost closed and the differential pressure over the valve is low, there is a nonlinear behavior where the volume flow rate sharply cuts from 0.3 to 0, which is clearly not captured in the model. Valve 4 is the valve which is the furthest from the pump and hence experiences the lowest differential pressure. Hence the volume flow rate in the other valves never becomes sufficiently low to exhibit this behavior. Finally there are outliers in the prediction errors e_3. These outliers are likely caused by measurement errors which can occur due to e.g. air bubbles forming in the system.

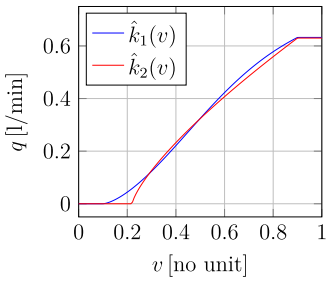
One difference between models B and C is that the prediction errors for model B appear segmented in one upper and one lower cluster. This is most clearly visible in valves 1 and 2. In Fig. 12 we show an enlarged view of these plots. Here we highlight the data points where the change in valve set-point is positive in red ($v_i(t) - v_i(t-1) > 0$), and negative in blue. The two clusters are clearly separated by the movement direction of the valve set-point, indicating the effect of hysteresis. In model C, this

Table 1
Pipe parameters estimated for each model structure on either the exciting (E) or realistic (R) dataset.

Model	Data	s_1	s_2	s_3	s_4	s_5	s_6	s_7
A	E	0.0	0.0	0.0	0.0	0.0089	0.00082	0.021
B	E	0.0	0.0	0.0	0.00067	0.0039	0.0046	0.029
C	E	0.0	0.0	0.0	0.015	0.0038	0.0045	0.029
C	R	0.0	0.0	0.0	0.0	0.0044	0.0	0.049



(a) Hydraulic schematic of one consumer in the district heating network, represented by two symmetrical pipes with combined resistance s and valve curve $K_v k(v)$.



(b) Plots of $\hat{k}_1(v)$ and $\hat{k}_2(v)$ which relate the flow rate q and the valve position v when $p_i - p_j = 1$.

Fig. 8. An illustration of a consumer connected to the network, represented by two symmetrical pipes and a valve. For a large portion of the valve curve, two different models of the hydraulic admittance of this consumer could have very similar properties. Here $\hat{k}_1(v)$ represents a pipe and a valve, but $\hat{k}_2(v)$ represents only a valve with no pipe. The difference between the two models becomes evident only when the valve is almost closed.

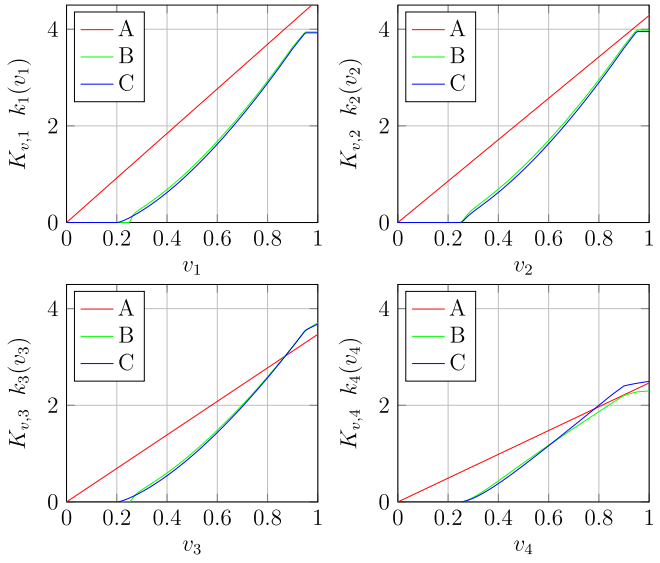


Fig. 9. Valve curves $K_{v,i}k_i(v_i)$ for each valve i under each model structure A, B and C trained on the exciting data set.

effect is no longer visible, demonstrating the strength of the hysteresis compensation.

In total, model C is the strongest model when trained and tested on the exciting data set. Histograms of the prediction errors $e_i(t)$ for this model are shown in Fig. 13. All four error distributions have large tails. In Fig. 10 we saw that these tails arise due to not fully capturing the valve characteristics for valves 1 and 2 and the data outliers of valves 3 and 4. Outside of these tails, a vast majority of the errors lie in the range $[-0.2, 0.2]$l/min, which can be compared with the mean flow rates $\bar{q}_i$ exhibited in the system, given by $\bar{q}_1 = 4.60$ l/min, $\bar{q}_2 = 4.34$ l/min, $\bar{q}_3 = 2.91$ l/min, and $\bar{q}_4 = 2.11$ l/min respectively in the exciting data set. Hence in the well-tuned regions of the valve curve, errors remain within 5 or 10% of the mean flow rates for valves 1–2 and 3–4 respectively.

5.2. Realistic training data

Fig. 14 shows the evaluated valve curves gained from training model structure C on the exciting or realistic data set respectively. We can note that in the lower regions of the valve curves, there is a significant overlap between the two curves. However, above a valve set-point of roughly 0.6, we see the curves starting to deviate. This is likely due to the fact that the model trained on the realistic data set has not seen any data within this region. This is also reflected in Fig. 15 which shows the prediction errors of using model C trained on the realistic data set. The model behaves rather well when evaluated on the realistic data set and maintains the errors within about 0.3 l/min. However, the realistic data only represents valve set-points in limited intervals. The green bar in the figures highlights the interval between the 5th and 95th quantiles of the valve set-points in the training data. The model is unable to extrapolate beyond the limited intervals of valve set-points represented in the training data which is reflected in the plots on the right of Fig. 15. Here the model is evaluated on the exciting data set, which covers a larger range of valve operation. In the region of the valve set-points where the model was trained, prediction errors are relatively small. Outside of these intervals however, the models behave extremely poorly. This is not surprising, as no such data was used in training. While it may be difficult to guarantee that the full operational range of each valve is represented in data from a real system, it may also be unlikely that the operational range present in the data is quite so restrictive as in our example. Firstly, if such a small range of a valves operational ranges is used, this would imply that the valve may be poorly dimensioned for the building in which it is installed. Secondly, we are only using two weeks worth of representational data. We can see already in Fig. 15 that the last 30% of the data has been shifted in comparison to the training data. Hence if a longer period of time is considered, it becomes increasingly likely that a larger portion of the valves operational range is explored. Thirdly, the reference values for this data are based on hourly means of recorded volume flow rates. The mean operation naturally hides the peaks and valleys of the operational modes and hence one could expect larger variance in a real data set.

6. Conclusion

In this paper we modeled the hydraulic properties of a district heating network, and tuned the parameters of said model based on operational data. This method was evaluated on laboratory data from the Smart Water Infrastructure laboratory in Aalborg. We found that in a realistic setting, the characteristics of valves are hugely important for accurately describing the system properties. A simple assumption of linear valve characteristics is therefore a naive approach. To remedy this issue, we proposed modeling the valve characteristic with a linear combination of many possible valve characteristics, which improved model performance. Additionally, we found that valve hysteresis played a significant role in the setup, which we remedied through the use of a preprocessing filter before doing model calibration. We compared the modeling procedure on two data set, one designed to test the system in all of its' possible operational mode and one designed to mimic an operational district heating network. We found that in the more

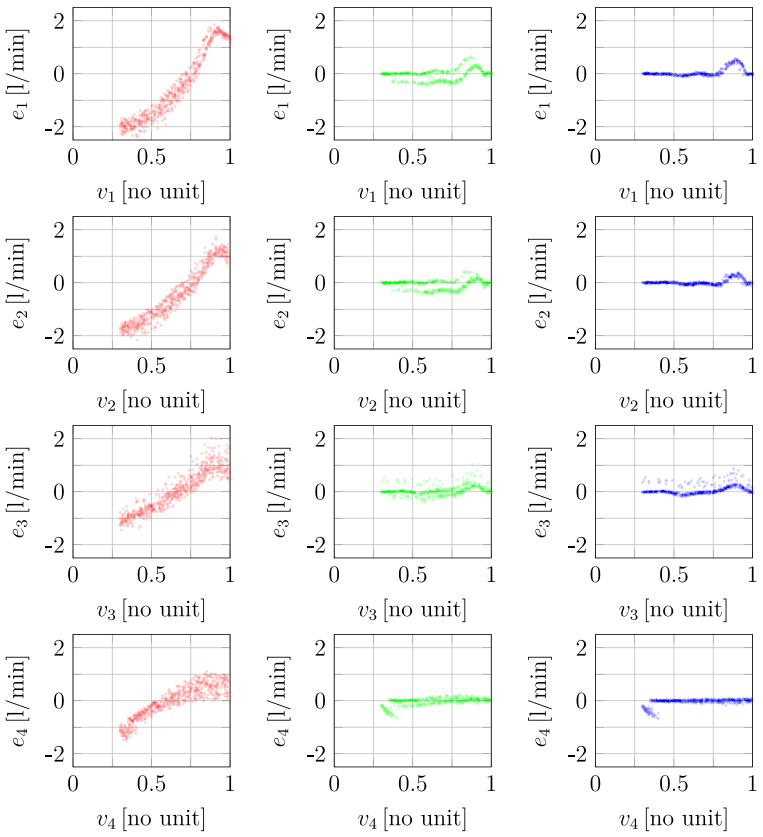


Fig. 10. Prediction errors $e_i = q_i - \hat{q}_i$ plotted over the valve set-point v_i for each valve i and each model structure A, B and C trained on the exciting data set. The plots are arranged as valve 1–4 from top to bottom and model A (red), B (green), and C (blue) from left to right.

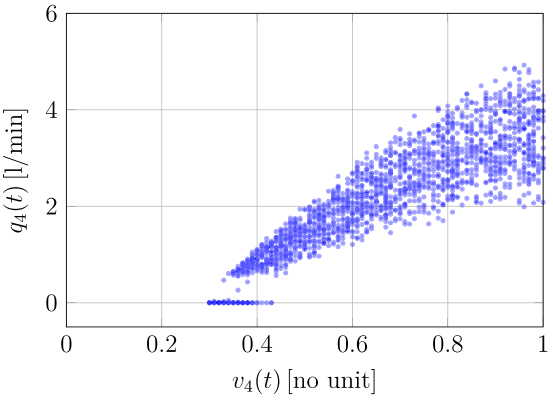


Fig. 11. Scatter-plot of flow rates $q_4(t)$ against valve set-points $v_4(t)$ in the exciting data set.

realistic data set, the valves only operated in a limited region. This led to the model extrapolating poorly to operational modes outside of this region. However, in the first data set where a larger portion of the systems operational range was explored, the best performing model was able to predict volume flow rates within 5 and 10% deviation from the mean values, excluding the portion of the valve curve where the fit was the most poor. The performance in this region could likely be improved by further tailoring the parameterization of the valve curve.

7. Future work

There are several open questions and directions for extending the results of this paper. Firstly, it is not common in current district heating networks to have access to measurements of valve positions. While this data may become available with future smart installations, a current approach could be to jointly estimate the model parameters and the valve positions. Such an approach would present many interesting challenges, as the quality of the parameter estimation depends on the accuracy of the valve position estimation, and the valve position estimation conversely depends on the accuracy of the model calibration.

A further investigation can be conducted with regards to choosing the filtering parameter δ for hysteresis compensation. In this experimental setup we have only four valves which can be assumed to have similar characteristics. Thus a shared, heuristically chosen δ is viable. For a larger and more heterogeneous network, this approach is no longer valid and thus a reasonable search approach is needed.

It appears that a further investigation into the exponent γ relating volume flow rates and differential pressures over pipes and valves could be an avenue to model improvement. However, such an investigation can quickly become complex. E.g. should this exponent be shared between valves and pipes, and should each component have its own exponent? Tuning this parameter based on data is a non-trivial task with regards to maintaining computational tractability.

Finally one can consider an extension to meshed distribution networks. This causes the issue that the volume flow rates in pipes in the supply-and-return networks are not explicitly known by measuring the valve volume flow rates. Previous works have solved for these values in tandem with model parameters using non-convex optimizers such as particle swarm solvers [10]. Such previous works have however been limited to considering only pipe resistances and not valve characteristics. Therefore they could be less concerned with the computational complexity and convergence issues which may differentiate non-convex from convex problems.

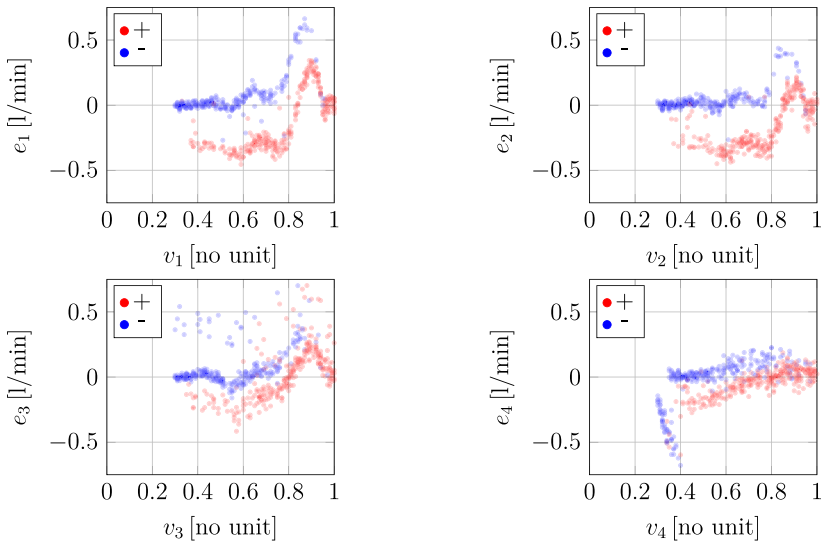


Fig. 12. Prediction errors $e_i = q_i - \hat{q}_i$ for each valve i using model B trained on the exciting data set, scattered over the valve set-points v_i. The points where the valve opened since the last data point ($v_i(t) - v_i(t-1) > 0$) are shown in red, and the points where the valve closed are shown in blue.

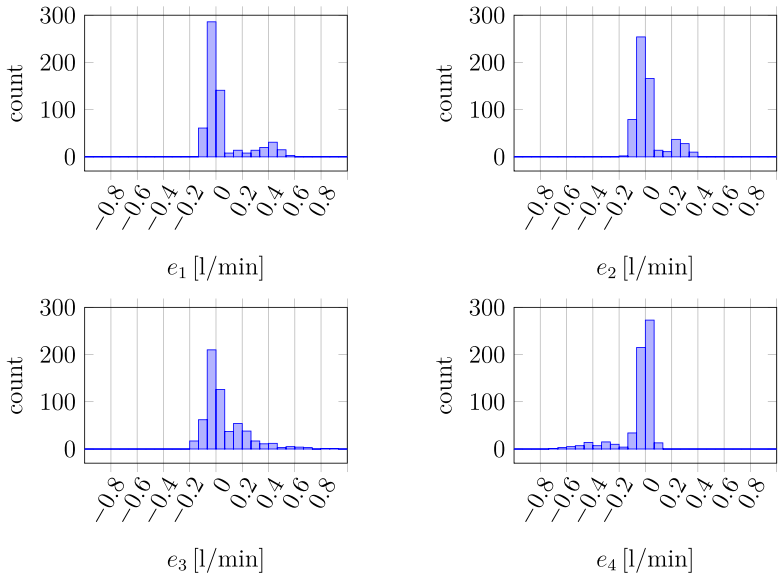


Fig. 13. Histograms of prediction errors $e_i = q_i - \hat{q}_i$ using model structure C trained on the exciting data set.

CRediT authorship contribution statement

Felix Agner: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Christian Møller Jensen:** Writing – original draft, Visualization, Software. **Anders Rantzer:** Writing – review & editing, Supervision. **Carsten Skovmose Kallesøe:** Writing – review & editing, Conceptualization. **Rafal Wisniewski:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Felix Agner and Anders Rantzer are members of the ELLIIT Strategic Research Area at Lund University.

This work is funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement No 834142 (ScalableControl).

The work has been partially supported by the Poul Due Jensen Foundation, Denmark through the Smart Water Infrastructure Laboratory project and the Swift project.

Christian Møller Jensen is funded by Danish Innovation Fund, Denmark grant no. 3129-00019B.

Appendix A. Forward estimation

In a line-structured network such as the one we investigate, our parameterized model can be used for forward estimation. This implies calculating the resulting flow rates q_i given the valve set-points v_i and the differential pressure Δp_0. To accomplish this we can realize that given a parameterized valve model on the form (4) and given a valve set-point v_i, (4) can be replaced by

$$\Delta p_i = r_i q_i^2 \tag{A.1}$$

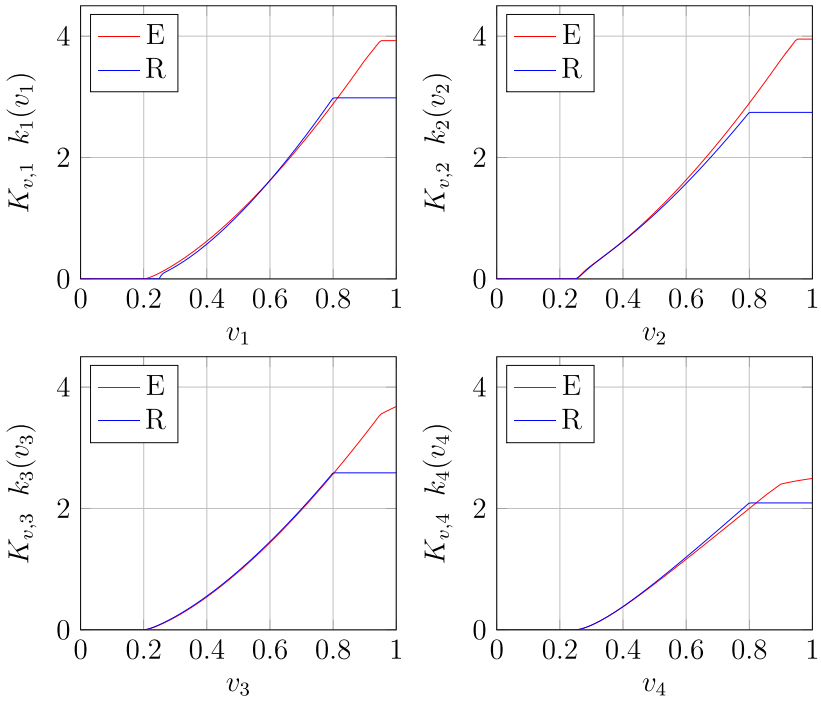


Fig. 14. Valve curves $K_{v,i}k_i(v_i)$ for each valve i for model structure C trained on the exciting (E, red) and realistic (R, blue) datasets respectively.

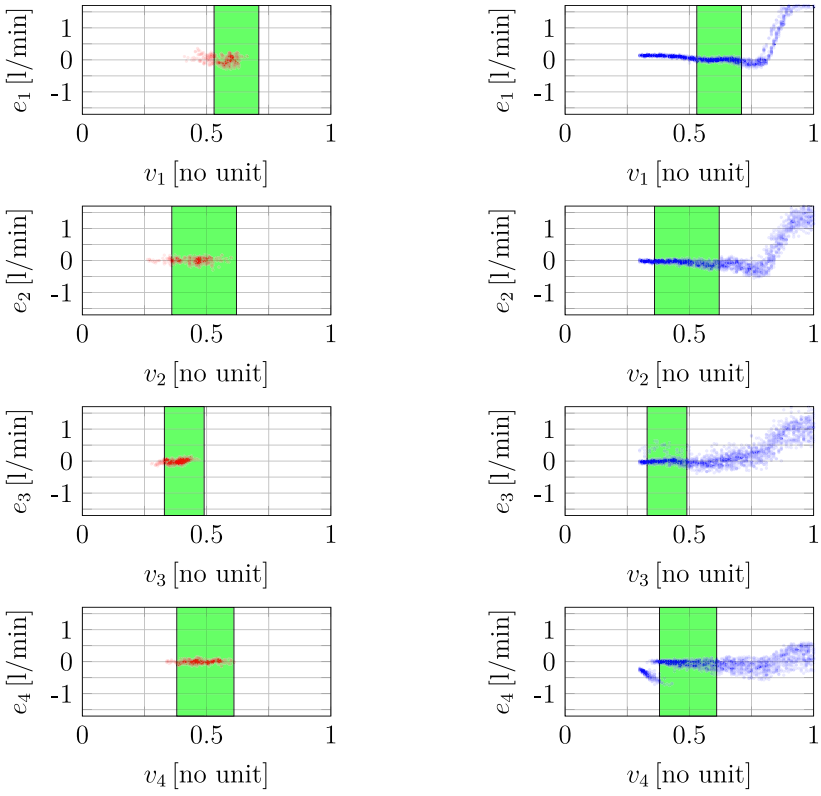
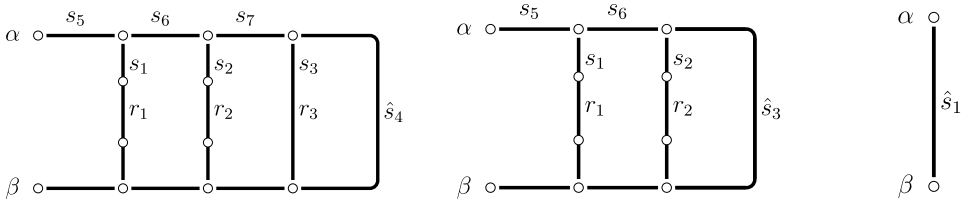


Fig. 15. Prediction errors $e_i = q_i - \hat{q}_i$ for each valve i using model C trained on the realistic data set and evaluated on both the realistic data set (red, left) and the exciting data set (blue, right). The figures are arranged as valve 1 to 4 from top to bottom. The green bar outlines the interval between the 5th and 95th quantiles for v_i in the realistic training data.



(a) Valve resistance r_4 and pipe resistances s_4 are replaced by one equivalent resistance $\hat{s}_4$.

(b) The parallel connection of valves 3 and 4 are reduced to one equivalent resistance $\hat{s}_3$.

(c) The network is reduced to one equivalent resistance $\hat{s}_1$.

Fig. A.1. Step-by-step illustration for reducing the full hydraulic representation of the network into one equivalent resistance with given valve set-points.

where $r_i = \left(\frac{\theta_{i,1}}{k_1(v_i)^2} + \frac{\theta_{i,2}}{k_2(v_i)^2} + \cdots + \frac{\theta_{i,K}}{k_K(v_i)^2}\right)$. In this way once the spindle position is fixed, a valve functions like a pipe in the sense of how it relates pressure drops and flow rates. To calculate the total amount of water flowing from α to β, we can reduce our system in a series of steps to one single, equivalent resistance (see Fig. A.1). First we replace s_4 and r_4 with an equivalent resistance $\hat{s}_4 = 2s_4 + r_4$ (see Fig. A.1(a)). We can now find a relation between q_3 and q_4 by establishing that

$$\hat{s}_4 q_4^2 = (2s_3 + r_3)q_3^2$$

and hence

$$q_3 = \sqrt{\frac{\hat{s}_4}{2s_3 + r_3}}\, q_4.$$

We will now lump the part of the network right of valve 2 into one equivalent resistance $\hat{s}_3$ (see Fig. A.1(b)). The flow rate going through $\hat{s}_3$ is $q_3 + q_4$, and hence it must hold that

$$\hat{s}_3(q_3+q_4)^2 = 2s_6\left(q_3+q_4\right)^2 + \hat{s}_4 q_4^2 \tag{A.2}$$

$$\begin{aligned}\implies \hat{s}_3 &= 2s_6 + \frac{q_4^2}{(q_3+q_4)^2}\hat{s}_4 \\ &= 2s_6 + \frac{q_4^2}{q_4^2\left(\sqrt{\hat{s}_4/(2s_3+r_3)}+1\right)^2}\hat{s}_4 \\ &= 2s_6 + \frac{\left(2s_3+r_3\right)\hat{s}_4}{\sqrt{\hat{s}_4}+\sqrt{2s_3+r_3}}.\end{aligned} \tag{A.3}$$

We can iteratively calculate $\hat{s}_i$ in this way according to

$$\hat{s}_i = 2s_{i+3} + \frac{\left(2s_i + r_i\right)\hat{s}_{i+1}}{\sqrt{\hat{s}_{i+1}} + \sqrt{2s_i + r_i}} \tag{A.4}$$

until reaching $\hat{s}_1$ which directly connects α to β (see Fig. A.1(c)). Now clearly the total flow rate $\hat{q}_1 = q_1 + q_2 + q_3 + q_4$ must satisfy

$$\hat{q}_1 = \sqrt{\frac{\Delta p_0}{\hat{s}_1}}. \tag{A.5}$$

Now we can calculate the individual flow rates q_1, q_2, q_3 and q_4 by expanding the network again in the other direction. We here note that it must hold that

$$\begin{aligned}&\hat{s}_1\hat{q}_1^2 = \hat{s}_2\left(\hat{q}_1 - q_1\right)^2 \\ \implies\ & q_1 = \left(1 - \sqrt{\frac{\hat{s}_1}{\hat{s}_2}}\right)\hat{q}_1\end{aligned}$$

and also that $\hat{q}_2 = q_2 + q_3 + q_4 = \hat{q}_1 - q_1$. We can thus calculate all of the flow rates recursively according to the rules

$$q_i = \left(1 - \sqrt{\frac{\hat{s}_i}{\hat{s}_{i+1}}}\right)\hat{q}_i \tag{A.6}$$

$$\hat{q}_{i+1} = \hat{q}_i - q_i. \tag{A.7}$$

Appendix B. Valve characteristic equations

The expressions relating flow rates, valve set-points and differential pressure for the valves in each model can be written down on closed form. For model A trained on the exciting data set, this relation is given by

$$\Delta p_1 = \frac{0.047}{v_1^2} q_1^2 \tag{B.1a}$$

$$\Delta p_2 = \frac{0.054}{v_2^2} q_2^2 \tag{B.1b}$$

$$\Delta p_3 = \frac{0.083}{v_3^2} q_3^2 \tag{B.1c}$$

$$\Delta p_4 = \frac{0.16}{v_4^2} q_4^2. \tag{B.1d}$$

For model B trained on the exciting data set, the expression is given by

$$\begin{aligned}\Delta p_1 = \Bigg(&\frac{0.011}{\mathrm{ramp}_{0.15}^{0.9}\left(v_1\right)^{3.0}} + \frac{0.049}{\mathrm{ramp}_{0.15}^{0.95}\left(v_1\right)^{3.0}} + \frac{0.0041}{\mathrm{ramp}_{0.2}^{0.95}\left(v_1\right)^{3.0}} \\ &+ \frac{0.00015}{\mathrm{ramp}_{0.25}^{0.95}\left(v_1\right)^{3.0}}\Bigg) q_1^2\end{aligned} \tag{B.2a}$$

$$\begin{aligned}\Delta p_2 = \Bigg(&\frac{0.009}{\mathrm{ramp}_{0.1}^{0.9}\left(v_2\right)^{3.0}} + \frac{0.0068}{\mathrm{ramp}_{0.1}^{0.95}\left(v_2\right)^{3.0}} + \frac{0.039}{\mathrm{ramp}_{0.15}^{0.95}\left(v_2\right)^{3.0}} \\ &+ \frac{0.0068}{\mathrm{ramp}_{0.2}^{0.95}\left(v_2\right)^{3.0}} + \frac{0.0011}{\mathrm{ramp}_{0.25}^{0.95}\left(v_2\right)^{3.0}}\Bigg) q_2^2\end{aligned} \tag{B.2b}$$

$$\begin{aligned}\Delta p_3 = \Bigg(&\frac{0.039}{\mathrm{ramp}_{0.15}^{0.95}\left(v_3\right)^{3.0}} + \frac{0.029}{\mathrm{ramp}_{0.15}^{1.0}\left(v_3\right)^{3.0}} + \frac{0.0044}{\mathrm{ramp}_{0.2}^{1.0}\left(v_3\right)^{3.0}} \\ &+ \frac{0.00072}{\mathrm{ramp}_{0.25}^{1.0}\left(v_3\right)^{3.0}}\Bigg) q_3^2\end{aligned} \tag{B.2c}$$

$$\begin{aligned}\Delta p_4 = \Bigg(&\frac{0.11}{\mathrm{ramp}_{0.1}^{0.9}\left(v_4\right)^{2.0}} + \frac{0.031}{\mathrm{ramp}_{0.1}^{1.0}\left(v_4\right)^{2.0}} + \frac{0.0043}{\mathrm{ramp}_{0.2}^{0.9}\left(v_4\right)^{3.0}} \\ &+ \frac{0.0078}{\mathrm{ramp}_{0.25}^{0.9}\left(v_4\right)^{3.0}} + \frac{0.033}{\mathrm{ramp}_{0.25}^{0.95}\left(v_4\right)^{3.0}}\Bigg) q_4^2\end{aligned} \tag{B.2d}$$

For model C trained on the exciting data set, the expressions are

$$\Delta p_1 = \left(\frac{0.0093}{\mathrm{ramp}_{0.15}^{0.9}\left(v_1\right)^{3.0}} + \frac{0.03}{\mathrm{ramp}_{0.15}^{0.95}\left(v_1\right)^{3.0}} + \frac{0.025}{\mathrm{ramp}_{0.2}^{0.95}\left(v_1\right)^{3.0}}\right) q_1^2 \tag{B.3a}$$

$$\Delta p_2 = \left(\frac{0.0065}{\mathrm{ramp}_{0.15}^{0.9}\left(v_2\right)^{3.0}} + \frac{0.038}{\mathrm{ramp}_{0.15}^{0.95}\left(v_2\right)^{3.0}} + \frac{0.018}{\mathrm{ramp}_{0.2}^{0.95}\left(v_2\right)^{3.0}}\right.$$

$$+ \frac{0.0022}{\mathrm{ramp}_{0.25}^{0.95}\left(v_2\right)^{3.0}}\Bigg) q_2^2 \tag{B.3b}$$

$$\Delta p_3 = \left(\frac{0.04}{\mathrm{ramp}_{0.15}^{0.95}\left(v_3\right)^{3.0}} + \frac{0.0089}{\mathrm{ramp}_{0.2}^{0.95}\left(v_3\right)^{3.0}} + \frac{0.025}{\mathrm{ramp}_{0.2}^{1.0}\left(v_3\right)^{3.0}}\right) q_3^2 \tag{B.3c}$$

$$\Delta p_4 = \Bigg(\frac{0.023}{\mathrm{ramp}_{0.1}^{0.85}\left(v_4\right)^{2.0}} + \frac{0.069}{\mathrm{ramp}_{0.1}^{0.9}\left(v_4\right)^{2.0}} + \frac{0.0085}{\mathrm{ramp}_{0.2}^{0.9}\left(v_4\right)^{3.0}} + \frac{0.036}{\mathrm{ramp}_{0.25}^{0.9}\left(v_4\right)^{3.0}} + \frac{0.001}{\mathrm{ramp}_{0.25}^{0.95}\left(v_4\right)^{3.0}} + \frac{0.023}{\mathrm{ramp}_{0.25}^{1.0}\left(v_4\right)^{3.0}}\Bigg) q_4^2 \tag{B.3d}$$

Finally, when trained on the realistic data set, model C yields the following expressions.

$$\Delta p_1 = \left(\frac{0.11}{\mathrm{ramp}_{0.2}^{0.8}\left(v_1\right)^{3.0}} + \frac{0.00014}{\mathrm{ramp}_{0.25}^{0.8}\left(v_1\right)^{3.0}}\right) q_1^2 \tag{B.4a}$$

$$\Delta p_2 = \left(\frac{0.11}{\mathrm{ramp}_{0.15}^{0.8}\left(v_2\right)^{3.0}} + \frac{0.024}{\mathrm{ramp}_{0.25}^{0.8}\left(v_2\right)^{2.5}} + \frac{0.0029}{\mathrm{ramp}_{0.25}^{0.8}\left(v_2\right)^{3.0}}\right) q_2^2 \tag{B.4b}$$

$$\Delta p_3 = \left(\frac{0.085}{\mathrm{ramp}_{0.15}^{0.8}\left(v_3\right)^{3.0}} + \frac{0.065}{\mathrm{ramp}_{0.2}^{0.8}\left(v_3\right)^{3.0}}\right) q_3^2 \tag{B.4c}$$

$$\Delta p_4 = \left(\frac{0.077}{\mathrm{ramp}_{0.1}^{0.8}\left(v_4\right)^{2.0}} + \frac{0.026}{\mathrm{ramp}_{0.1}^{0.8}\left(v_4\right)^{2.5}} + \frac{0.12}{\mathrm{ramp}_{0.25}^{0.8}\left(v_4\right)^{3.0}}\right) q_4^2. \tag{B.4d}$$

Data availability

The data sets prepared and used in this paper are openly available via GitHub under a CC-BY 4.0 license [25].

References

[1] Lund H, Werner S, Wiltshire R, Svendsen S, Thorsen JE, Hvelplund F, et al. 4Th generation district heating (4GDH). Energy 2014;68:1–11. http://dx.doi.org/10.1016/j.energy.2014.02.089.

[2] Vandermeulen A, van der Heijde B, Helsen L. Controlling district heating and cooling networks to unlock flexibility: A review. Energy 2018;151:103–15. http://dx.doi.org/10.1016/j.energy.2018.03.034.

[3] Benonysson A, Bøhm B, Ravn HF. Operational optimization in a district heating system. Energy Convers Manage 1995;36(5):297–314. http://dx.doi.org/10.1016/0196-8904(95)98895-T.

[4] Vesterlund M, Toffolo A, Dahl J. Simulation and analysis of a meshed district heating network. Energy Convers Manage 2016;122:63–73. http://dx.doi.org/10.1016/j.enconman.2016.05.060.

[5] Wang Y, Shi K, Zheng X, You S, Zhang H, Zhu C, et al. Thermo-hydraulic coupled analysis of meshed district heating networks based on improved breadth first search method. Energy 2020;205:117950. http://dx.doi.org/10.1016/j.energy.2020.117950.

[6] Dénarié A, Aprile M, Motta M. Dynamical modelling and experimental validation of a fast and accurate district heating thermo-hydraulic modular simulation tool. Energy 2023;282:128397. http://dx.doi.org/10.1016/j.energy.2023.128397.

[7] Guelpa E, Toro C, Sciacovelli A, Melli R, Sciubba E, Verda V. Optimal operation of large district heating networks through fast fluid-dynamic simulation. Energy 2016;102:586–95. http://dx.doi.org/10.1016/j.energy.2016.02.058.

[8] Wang Y, You S, Zhang H, Zheng W, Zheng X, Miao Q. Hydraulic performance optimization of meshed district heating network with multiple heat sources. Energy 2017;126:603–21. http://dx.doi.org/10.1016/j.energy.2017.03.044.

[9] Wang N, You S, Wang Y, Zhang H, Miao Q, Zheng X, et al. Hydraulic resistance identification and optimal pressure control of district heating network. Energy Build 2018;170:83–94. http://dx.doi.org/10.1016/j.enbuild.2018.04.003.

[10] Zheng X, Shi Z, Wang Y, Zhang H, Liu H. Thermo-hydraulic condition optimization of large-scale complex district heating network: A case study of Tianjin. Energy 2023;266:126406. http://dx.doi.org/10.1016/j.energy.2022.126406.

[11] Agner F, Kergus P, Pates R, Rantzer A. Combating district heating bottlenecks using load control. Smart Energy 2022;6:100067. http://dx.doi.org/10.1016/j.segy.2022.100067.

[12] Bahlawan H, Ferraro N, Gambarotta A, Losi E, Manservigi L, Morini M, et al. Detection and identification of faults in a district heating network. Energy Convers Manage 2022;266:115837. http://dx.doi.org/10.1016/j.enconman.2022.115837.

[13] Savic DA, Kapelan ZS, Jonkergouw PM. Quo vadis water distribution model calibration? Urban Water J 2009;6(1):3–22. http://dx.doi.org/10.1080/15730620802613380.

[14] Liu Y, Wang P, Luo P. Pipe hydraulic resistances identification of district heating networks based on matrix analysis. Energies 2020;13(11). http://dx.doi.org/10.3390/en13113007.

[15] Luo P, Wang H, Liu Y, Du Q, Zhang Z. Resistance characteristic parameters estimation of hydraulic model in heating networks based on real-time operation data. Buildings 2022;12(6). http://dx.doi.org/10.3390/buildings12060743.

[16] Agner F, Kergus P, Pates R, Rantzer A. Hydraulic parameter estimation in district heating networks. IFAC-PapersOnLine 2023;56(2):5438–43. http://dx.doi.org/10.1016/j.ifacol.2023.10.194, 22nd IFAC World Congress.

[17] Zheng X, Shi Z, Wang Y, Zhang H, Tang Z. Digital twin modeling for district heating network based on hydraulic resistance identification and heat load prediction. Energy 2024;288:129726. http://dx.doi.org/10.1016/j.energy.2023.129726.

[18] Hägglund T. Process control in practice. Berlin, Boston: De Gruyter; 2023, http://dx.doi.org/10.1515/9783111104959.

[19] Val Ledesma J, Wisniewski R, Kallesøe CS. Smart water infrastructures laboratory: Reconfigurable test-beds for research in water infrastructures management. Water 2021;13(13). http://dx.doi.org/10.3390/w13131875.

[20] De Persis C, Kallesoe CS. Pressure regulation in nonlinear hydraulic networks by positive and quantized controls. IEEE Trans Control Syst Technol 2011;19(6):1371–83. http://dx.doi.org/10.1109/TCST.2010.2094619.

[21] Frederiksen S, Werner S. District heating and cooling. Studentlitteratur; 2013.

[22] Diamond S, Boyd S. CVXPY: A Python-embedded modeling language for convex optimization. J Mach Learn Res 2016;17(83):1–5.

[23] Agrawal A, Verschueren R, Diamond S, Boyd S. A rewriting system for convex optimization problems. J Control Decision 2018;5(1):42–60.

[24] Vallabha G. Real-time pacer for simulink. 2016, Published in Matlab Central repository.

[25] Agner F. DH hydraulic parameter estimation. 2024, http://dx.doi.org/10.5281/zenodo.11946409.