Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2021) 178-193

ECCOMAS
Proceedia

UNCECOMP 2021
4th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Streamed from Athens, Greece, 28 -30 June 2021

OPTIMAL SENSOR PLACEMENT IN DISTRICT HEATING NETWORKS FOR BAYESIAN INFERENCE OF UNCERTAIN DEMANDS

A. Matei[1], A. Bott[2], L. Rehlich[1], F. Steinke[2] and S. Ulbrich[1]

[1] Technical University of Darmstadt, Department of Mathematics, Research Group Optimization
Dolivostraße 15, 64293 Darmstadt, Germany
e-mail for correspondence: matei@mathematik.tu-darmstadt.de

[2] Technical University of Darmstadt, Department of Electrical Engineering and Information Technology, Energy Information Networks and Systems Lab
Landgraf-Georg Straße 4, 64289, Darmstadt, Germany
e-mail for correspondence: andreas.bott@eins.tu-darmstadt.de

Abstract. *District heating networks have traditionally been designed and operated as distribution networks supplied by few central heat production units. In order to reduce emissions in the heating sector feed-in from smaller decentralized units and industrial waste heat is becoming ever more important. This transition requires a more detailed monitoring of the network state, which can be achieved either by installing a huge number of additional sensors in the grid or by a state estimation based on a few additional measurements. However, the uncertain heat consumption generated by consumers presents a major challenge in this endeavor. In this paper we propose a model-based approach for optimal sensor placement in district heating networks in order to minimize the uncertainty in the demand values which are estimated by solving a Bayesian inverse problem. The optimization scheme is designed to yield a fair compromise between the desired information gain and the costs for installing the chosen sensors. A steady-state model is employed to estimate temperatures, mass flows and pressures of network components given the mean demand and the initial pressure generated by the heating plant. Our approach is applied to a real-sized district heating network using actual consumption distributions as given prior in order to validate the model and to prove scalability.*

Keywords: heating networks, sensor placement, Bayesian inference, demand estimation

ISSN:2623-3339
Peer-review under responsibility of the organizing committee of UNCECOMP 2021.
doi: 10.7712/120221.8031.19135

1 INTRODUCTION

District heating networks are closed looping systems in the sense that water is pumped from heating plants towards consumers through a pipe system and flows back through parallel laid pipes. Energy is transmitted by heating up the water at the heating plants and cooling it down at the consumer's place. Within the grid the energy flow through each element is not predetermined by the network operator but is a result of the energy extracted by the consumers. In order to ensure security of supply and minimize energy losses in the grid, temperature and pressure at the heating plant have to be adjusted accordingly.

Future district heating networks will be characterized by lower temperature levels and additional decentralized feed-in [16] making it considerably more difficult to ensure, that no grid element is overloaded. Therefore additional information about the network state and consequently the consumer's actual consumption have to be obtained. However, constantly measuring the demand at each consumer's place is not practically feasible. The large number of additional sensors would not only mean high investment cost but also a high onside electricity demand. Moreover measurements directly at the consumers may be misleading, due to the way they are connected to the grid. Usually the main pipes are laid under the roads and smaller pipes connect these with the heat-exchange-stations inside the buildings. If the consumption changes drastically or is close to zero the network state inside this connection pipes is not representative for the main pipes in the grid.

Alternatively, the heat consumption can be predicted or estimated based on external parameters [6]. These estimations naturally inherit some kind of uncertainty affecting the model prediction [18]. Hence, we say that uncertainty propagates from the consumption parameters to the network's state via the parameter-to-observable map. In this paper we propose a model-based approach to place a small number of sensors at optimal positions in the district heating network in order to minimize the variance of the estimated demand values. We claim that a better knowledge about the consumer's consumption eventually leads to more precise state predictions in the whole network. In order to quantify the uncertainty of both the demand estimation and the model's prediction we use the linearized parameter-to-observable map and a Bayesian viewpoint.

Optimal sensor placement is a broad field of research. It often appears in the context of optimal experimental design in the literature [2, 17, 10, 12, 21]. Probabilistic sensitivity-based approaches [3, 14, 19] and Bayesian inference-based perspectives [1, 2, 11] are mainly used as a tool to maximize the information gain obtained from optimally positioned sensors at low cost. A topic that is closely related to our question is leakage detection. A review paper on leakage detection methods in district heating networks is given by [24]. In [5] a distributed demand response approach based on augmented Lagrangian methods to optimize the heating demand with minimal private information exchange is developed. To the best of our knowledge, the sensor placement problem for variance-minimal demand estimation has not been applied to district heating networks so far.

The paper is structured as follows. In Section 2 we introduce a heating model which maps the consumer's demand onto the network's state. A Bayesian inference approach for model-based optimal experimental design is applied to our setting in Section 3. Numerical results for a real-sized district heating network are presented in Section 4. We end the paper with a short discussion of the results and a conclusion.

2 HEAT MODEL EQUATIONS

For practical purposes we are interested in the pressures and temperatures at given points in the network as well as the amount of water flowing through the pipes. Therefore, an operator $(\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\eta}) \mapsto \boldsymbol{e}(\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\eta})$ is constructed which couples the consumer demands $\boldsymbol{\theta}$, the network state $\boldsymbol{y}$ consisting of temperatures $\boldsymbol{T}$, pressures $\boldsymbol{p}$ and mass flows $\dot{\boldsymbol{m}}$, and set-point values $\boldsymbol{\eta}$. In the following we describe the different components of this operator $\boldsymbol{e}$. The structure of district heating networks can be described and implemented as a graph $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E})$ with nodes $\boldsymbol{V}$ and directed edges $\boldsymbol{E} \subseteq \boldsymbol{V} \times \boldsymbol{V}$. In this setting, the network state is defined according to the edges and nodes of the graph:

$$\boldsymbol{y} \coloneqq (p_i, T_i, T_{kl}^{\text{end}}, T_{kl}^{\text{start}}, \dot{m}_{kl}), \qquad \text{for all } i \in \boldsymbol{V} \text{ and } (k,l) \in \boldsymbol{E}, \tag{1}$$

where p_i, T_i are the pressure and the temperature in node i, and $T_{kl}^{\text{end}}, T_{kl}^{\text{start}}, \dot{m}_{kl}$ denote the water temperature at the end respectively the start of edge (k,l) and the mass flow on that edge. These kinds of models are commonly used to analyze different aspects of district heating networks [9, 15, 23, 4, 7]. In this paper, we assume steady state conditions and neglect time delays in the network. Since $\boldsymbol{G}$ is directed, we can assign a nominal flow direction to each edge. For an edge $(i,j) \in \boldsymbol{E}$ the nominal flow direction is from node i to node j, meaning that $\dot{m}_{ij} \geq 0$ if water flows form node i towards node j. The superscripts *start* and *end* should be understood in the sense of this nominal flow direction. Per definition it follows that $\dot{m}_{ij} = -\dot{m}_{ji}$ and $T_{ij}^{\text{end}} = T_{ji}^{\text{start}}$. We only investigate treelike networks with one heating plant, in which the mass flow directions cannot change. Therefore, we deliberately choose the nominal flow direction in such a way, that only positive values for the mass flow occur, i.e., if $(i,j) \in \boldsymbol{E}$ then $\dot{m}_{ij} \geq 0$.

Let $\boldsymbol{N}_i \coloneqq \{j \in \boldsymbol{V} \mid (i,j) \in \boldsymbol{E} \text{ or } (j,i) \in \boldsymbol{E}\}$ be the set of nodes in the neighborhood of i which are connected to node i by an edge $(i,j) \in \boldsymbol{E}$ or $(j,i) \in \boldsymbol{E}$. Furthermore, let

$$\boldsymbol{E}_i^+ \coloneqq \left\{(j,i) \in \boldsymbol{E} \mid j \in \boldsymbol{N}_i \text{ and } \dot{m}_{ji} > 0\right\} \quad \text{and} \quad \boldsymbol{E}_i^- \coloneqq \left\{(i,j) \in \boldsymbol{E} \mid j \in \boldsymbol{N}_i \text{ and } \dot{m}_{ij} \geq 0\right\}$$

denote the set of edges through which water flows into respectively out of node i. Kirchhoff's law can be applied to the network in the sense that the total mass of water which flows into a node, matches the total mass of water flowing out of a node:

$$\sum_{(j,i) \in \boldsymbol{E}_i^+} \dot{m}_{ji} = \sum_{(i,j) \in \boldsymbol{E}_i^-} \dot{m}_{ij}, \qquad \text{for all } i \in \boldsymbol{V}\,. \tag{2}$$

The temperature in each node is determined by the mixing laws of thermodynamics [23]:

$$T_i = \left(\sum_{(j,i) \in \boldsymbol{E}_i^+} \dot{m}_{ji}\, T_{ji}^{\text{end}} \right) \Bigg/ \left(\sum_{(j,i) \in \boldsymbol{E}_i^+} \dot{m}_{ji} \right), \qquad \text{for all } i \in \boldsymbol{V}\,. \tag{3}$$

Similarly, the temperature at the origin of an edge is given by the temperature of the node if the edge is an outflow of that node:

$$T_{ij}^{\text{start}} = T_i \qquad \text{if } (i,j) \in \boldsymbol{E}_i^-\,. \tag{4}$$

In our model, the four edge types $\boldsymbol{E}^{\text{load}}$, $\boldsymbol{E}^{\text{pipe}}$, $\boldsymbol{E}^{\text{heating}}$ and $\boldsymbol{E}^{\text{pump}}$ are distinguished. A single edge might represent multiple components in the physical network:

- The edges $\boldsymbol{E}^{\text{load}}$ represent consumers in the heating grid. In the physical world a heat exchanger is used to transfer energy from the district heating grid into the household heating system. A valve controls the mass flow through the heat exchanger in order to keep the water flowing back into the district heating network at a constant temperature. Additional equipment might be installed in order to measure the heat consumption or restrict the maximal mass flow [8]. This behavior is modeled as

$$T_{ij}^{\text{end}} = T_{ij}^{\text{set}}, \qquad \text{for all } (i,j) \in \boldsymbol{E}^{\text{load}} \tag{5}$$

$$\dot{m}_{ij} = \frac{\dot{Q}_{ij}}{c_p \left(T_{ij}^{\text{start}} - T_{ij}^{\text{end}}\right)}, \qquad \text{for all } (i,j) \in \boldsymbol{E}^{\text{load}} \tag{6}$$

$$p_i - p_j \geq \Delta p_{\min}, \qquad \text{for all } (i,j) \in \boldsymbol{E}^{\text{load}} \tag{7}$$

  where $c_{\text{p}} = 4.182\,\text{kJ}\,\text{kg}^{-1}\,\text{K}^{-1}$ is the specific heat capacity of water and T_{ij}^{set} and $\dot{Q}_{ij}$ are the consumer's set-points for the return temperature and the transferred heat energy, respectively. The model parameters $\boldsymbol{\theta}$ are exactly these heat energies $\dot{Q}_{ij}$ on a demand edge $(i,j) \in \boldsymbol{E}^{\text{load}}$. A minimum pressure difference $\Delta p_{\min}$ has to be applied by the grid in order to enable the flow through all components, cf. [8].

- Pipes $\boldsymbol{E}^{\text{pipe}}$ are passive elements in the grid, meaning that the change in pressure and temperature are not actively controlled but resulting from the mean mass flow through the pipe and the soil temperature T_{a}:

$$T_{ij}^{\text{end}} = \left(T_{ij}^{\text{start}} - T_{\text{a}}\right) \exp\left(-\frac{l_{ij}\lambda_{ij}}{c_p \dot{m}_{ij}}\right) + T_{\text{a}}\,, \qquad \text{for all } (i,j) \in \boldsymbol{E}^{\text{pipe}}, \tag{8}$$

$$p_i - p_j = \kappa_{ij} f_{\text{D},ij} \frac{8 l_{ij}}{\pi^2 \rho d_{ij}^5} \dot{m}_{ij}^2 + \rho g (z_j - z_i), \qquad \text{for all } (i,j) \in \boldsymbol{E}^{\text{pipe}}, \tag{9}$$

  compare [9, 23]. The coefficient λ_{ij} denotes the heat transferred through the isolation per pipe length and temperature difference between water and soil in $\text{W}\,\text{m}^{-1}\,\text{K}^{-1}$. Furthermore, $f_{\text{D},ij}$ is the Darcy friction factor which is can be calculated by the Colebrook-White equation

$$\frac{1}{\sqrt{f_{\text{D},ij}}} = -2\log_{10}\left(\frac{\epsilon_{ij}}{3.7 d_{ij}} + \frac{2.51}{\text{Re}_{ij}\sqrt{f_{\text{D},ij}}}\right) \tag{10}$$

  depending on the inner roughness ϵ_{ij} and the diameter d_{ij} of the pipes, as well as the Reynolds number Re_{ij}. The correction factor κ_{ij} is introduced in eq. (9) to account for the pressure loss due to bends. The constants κ_{ij}, d_{ij} and ϵ_{ij} are grid parameters that are assumed to be well known. Evidently, $\rho = 997\,\text{kg}\,\text{m}^{-3}$ is the density of water and $z_j - z_i$ is the difference of altitude between the nodes j and i.

- The heating edges $\boldsymbol{E}^{\text{heating}}$ are introduced to serve as slack edges to fulfill the law of energy conservation in the grid. Therefore the equations

$$T_{ij}^{\text{end}} = T_{ij}^{\text{set}} \qquad \text{for all } (i,j) \in \boldsymbol{E}^{\text{heating}} \tag{11}$$

$$p_j = p_i \qquad \text{for all } (i,j) \in \boldsymbol{E}^{\text{heating}}\,. \tag{12}$$

  restrict only the temperature at the end of an edge to be at a fixed set-point T_{ij}^{set} while the pressure does not change.

- Likewise, the pump edges $\boldsymbol{E}^{\text{pump}}$ are introduced as slack edges where the temperature

$$T_{ij}^{\text{end}} = T_{ij}^{\text{start}}, \qquad \text{for all } (i,j) \in \boldsymbol{E}_{ij}^{\text{pump}}, \tag{13}$$

  does not change. Typically, the model has only one heating and one pump edge which are directly connected in series and represent the largest heating plant in the network. The pressure difference between the supply side and the return side is at its highest at the heating plant and decreases with increasing distance. The demand edge $(v,w) \in \boldsymbol{E}^{\text{load}}$ with the lowest pressure difference is the so called worst point of the network. In real life networks this point is well known. The typical control scheme consists of measuring the pressure difference at this point and adjusting the pump pressure in such a way that the requirement in eq. (7) is just met for the worst point. Additionally an overall pressure level has to be maintained to prevent evaporation. These restrictions are mimicked in our model by fixing the pressures

$$p_v = p_0, \qquad p_w = p_0 - \Delta p_{\min}, \tag{14}$$

  with a suitable setpoint value p_0 measured in bar.

The functional relations (2) – (14) form a system of nonlinear equations which are summarized by the operator $\boldsymbol{e}(\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\eta})$ in a state equation

$$\boldsymbol{e}(\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\eta}) = 0, \tag{15}$$

which has a unique solution $\boldsymbol{y}(\boldsymbol{\theta}, \boldsymbol{\eta})$ for given consumer demands $\boldsymbol{\theta}$ and given set-point values $\boldsymbol{\eta} := (p_0, \Delta p_{\min}, p_{ij}^{\text{start}}, T_{ij}^{\text{set}})$ as introduced before. The proof of the uniqueness and existence of the solution can be done analogous to [9]. We solve eq. (15) by a Newton-method with projected gradients where the starting point is determined after a fixed point iteration according to [9, 23].

3 BAYESIAN INFERENCE AND OPTIMAL EXPERIMENTAL DESIGN

Model-based optimal design of experiments has the task to find a setup of experimental conditions, like sensor positions and control mechanisms, such that the model parameters can be estimated with minimal variance. In our setting we want to find optimal sensor positions in district heating networks such that the uncertainty in the estimated demand values $\boldsymbol{\theta}$ is minimized.

The model equation (15) brings the state $\boldsymbol{y}$, the demands $\boldsymbol{\theta}$ and the set-point values $\boldsymbol{\eta}$ into a functional relation. Our aim is to employ temperature-, pressure- and flow-sensors that measure the components of the solution $\boldsymbol{y}(\boldsymbol{\theta}, \boldsymbol{\eta})$. However, not all state variables in the network are of equal interest. We want to exclude the unreasonable sensor positions at the outset to reduce the dimension of the resulting optimization problem. Therefore, we select only a subset of the vector $\boldsymbol{y}$ as possible output channels that can be measured by sensors. Let $\boldsymbol{\Xi}$ be such a selection matrix. Thus, we introduce the overall parameter-to-observable map

$$\boldsymbol{\theta} \mapsto \boldsymbol{h}(\boldsymbol{\theta}) := \boldsymbol{\Xi} \cdot \boldsymbol{y}(\boldsymbol{\theta}, \boldsymbol{\eta}) \in \mathbb{R}^{n_s} \tag{16}$$

that maps the model parameters to n_s quantities of interest that can be directly measured by sensors. This mapping $\boldsymbol{h}(\boldsymbol{\theta})$ serves as our computer model that is commonly enhanced by a probabilistic point of view [22] where the collected data $\boldsymbol{z}$ is assumed to be subject to observational noise which is modeled as a random variable $\boldsymbol{\varepsilon}$:

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}\,. \tag{17}$$

Within a Bayesian framework, the posterior probability distribution of the estimated parameters is determined by the Bayes formula. Let $\gamma > 0$ and $\pi_0 \in \mathcal{N}(\boldsymbol{\theta}_0, \boldsymbol{\Gamma}_0/\gamma)$ be a Gaussian prior, capturing the a priori knowledge we have about $\boldsymbol{\theta}$. Furthermore, let $\boldsymbol{\varepsilon} \in \mathcal{N}(0, \boldsymbol{\Sigma})$ be a Gaussian random variable with density ρ and noise covariance matrix $\boldsymbol{\Sigma}$. We assume that the measurements obtained from different sensors are independently distributed and thus $\boldsymbol{\Sigma}$ is a diagonal matrix. Considering eq. (17), the data likelihood $\pi(z \mid \boldsymbol{\theta})$ has thus the density $\rho(z - \boldsymbol{h}(\boldsymbol{\theta}))$. Then the posterior $\pi(\boldsymbol{\theta} \mid z)$ is given by

$$\begin{aligned} \pi(\boldsymbol{\theta} \mid z) &\propto \pi(z \mid \boldsymbol{\theta})\pi_0 \\ &= \exp\left(-\frac{1}{2}\left\|z - \boldsymbol{h}(\boldsymbol{\theta})\right\|^2_{\boldsymbol{\Sigma}^{-1}} - \frac{\gamma}{2}\left\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\right\|^2_{\boldsymbol{\Gamma}_0^{-1}}\right), \end{aligned} \tag{18}$$

where $\|\boldsymbol{x}\|_{\boldsymbol{A}} := \sqrt{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}$ is the weighted norm of a vector $\boldsymbol{x}$ with a matrix $\boldsymbol{A}$. The maximum a posteriori estimator (MAP) is a point $\bar{\boldsymbol{\theta}}$ that maximizes this posterior probability distribution function:

$$\bar{\boldsymbol{\theta}}(z) := \operatorname{argmin}_{\boldsymbol{\theta}} \left(\frac{1}{2}\left\|z - \boldsymbol{h}(\boldsymbol{\theta})\right\|^2_{\boldsymbol{\Sigma}^{-1}} + \frac{\gamma}{2}\left\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\right\|^2_{\boldsymbol{\Gamma}_0^{-1}}\right), \tag{19}$$

compare [22].

In our case, the parameter-to-observable map is nonlinear and thus one cannot expect to obtain a posterior probability distribution that yields confidence regions which are analytically tractable. Therefore, we linearize the mapping $\boldsymbol{\theta} \mapsto \boldsymbol{h}(\boldsymbol{\theta})$ at the MAP point $\bar{\boldsymbol{\theta}}$ to obtain a Gaussian posterior whose covariance matrix is given by

$$\boldsymbol{C}_{\text{post}}\left(\bar{\boldsymbol{\theta}}\right) = \left(\boldsymbol{J}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{J} + \gamma \boldsymbol{\Gamma}_0^{-1}\right)^{-1}, \tag{20}$$

where $\boldsymbol{J}$ is the sensitivity matrix which is computed by the implicit function theorem:

$$\boldsymbol{J} := \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}} = -\boldsymbol{\Xi}\left[\frac{\partial \boldsymbol{e}\left(\boldsymbol{y}(\bar{\boldsymbol{\theta}}, \boldsymbol{\eta}), \bar{\boldsymbol{\theta}}, \boldsymbol{\eta}\right)}{\partial \boldsymbol{y}}\right]^{-1} \frac{\partial \boldsymbol{e}\left(\boldsymbol{y}(\bar{\boldsymbol{\theta}}, \boldsymbol{\eta}), \bar{\boldsymbol{\theta}}, \boldsymbol{\eta}\right)}{\partial \boldsymbol{\theta}} \tag{21}$$

The confidence region $G(\bar{\boldsymbol{\theta}}, \boldsymbol{C}_{\text{post}}, \alpha)$ around the MAP point $\bar{\boldsymbol{\theta}}$ to a level $1 - \alpha$, where $\alpha \in (0, 1)$, has then the analytical expression

$$G\left(\bar{\boldsymbol{\theta}}, \boldsymbol{C}_{\text{post}}, \alpha\right) := \left\{\boldsymbol{\theta} \in \mathbb{R}^{n_{\text{p}}} : (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})^\top \boldsymbol{C}_{\text{post}}^{-1}(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}) \leq \chi^2_{n_{\text{p}}}(1 - \alpha)\right\}, \tag{22}$$

where $\chi^2_{n_{\text{p}}}(1 - \alpha)$ is the quantile of the χ^2 distribution with n_{p} degrees of freedom, see [10].

We now introduce weights $\omega_k \in \{0, 1\}$ for each predefined sensor position $k = 1, \ldots, n_{\text{s}}$ such that $\omega_k = 1$ if, and only if, sensor k is used. Set $\boldsymbol{\Omega} := \operatorname{diag}(\omega_1, \ldots, \omega_{n_{\text{s}}})$ as the weight matrix containing the vector $\boldsymbol{\omega}$ on its diagonal. The knowledge received from the used sensors is added to the noise model $\boldsymbol{\varepsilon} \in \mathcal{N}(0, \boldsymbol{\Omega}^{-1}\boldsymbol{\Sigma})$, whereby a division by zero is set to infinity. This has the meaningful interpretation that an unused sensor yields an infinitely large covariance in the corresponding output channel, i.e., we know nothing about that quantity of interest.

These weights $\boldsymbol{\omega}$ directly influence the MAP point

$$\bar{\boldsymbol{\theta}}(z; \boldsymbol{\omega}) := \operatorname{argmin}_{\boldsymbol{\theta}} \left(\frac{1}{2}\left\|z - \boldsymbol{h}(\boldsymbol{\theta})\right\|^2_{\boldsymbol{\Omega}\,\boldsymbol{\Sigma}^{-1}} + \frac{\gamma}{2}\left\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\right\|^2_{\boldsymbol{\Gamma}_0^{-1}}\right), \tag{23}$$

the sensitivity $\boldsymbol{J}$ in eq. (21), the posterior covariance matrix of the parameters

$$\boldsymbol{C}_{\text{post}}\left(\bar{\boldsymbol{\theta}}(\boldsymbol{z};\boldsymbol{\omega}),\boldsymbol{\omega}\right) = \left(\boldsymbol{J}^{\top}\,\boldsymbol{\Omega}\,\boldsymbol{\Sigma}^{-1}\,\boldsymbol{J} + \gamma\,\boldsymbol{\Gamma}_0^{-1}\right)^{-1}, \tag{24}$$

as well as the confidence region in eq. (22), compare [2, 14].

In an optimally designed experiment, a fair compromise between the cost $\boldsymbol{c} \in \mathbb{R}_+^{n_\text{s}}$ of the used sensors and a measure Ψ of $\boldsymbol{C}_{\text{post}}$, representing the information gain by these sensors, is obtained. In order to reduce the computational complexity of the following problem and since $\boldsymbol{z}$ is difficult to obtain or not available beforehand, we set $\bar{\boldsymbol{\theta}}(\boldsymbol{z};\boldsymbol{\omega}) = \bar{\boldsymbol{\theta}}_0$ to an initial MAP estimation and keep it constant through the optimization. Let $\kappa > 0$ be a penalty factor and let $\|\cdot\|_0$ be the ℓ_0-"norm". We consider the optimal experimental design problem

$$\min_{\boldsymbol{\omega}\in\{0,1\}^{n_\text{s}}} \quad \Psi\left[\,\boldsymbol{C}_{\text{post}}\left(\bar{\boldsymbol{\theta}}_0,\boldsymbol{\omega}\right)\right] + \kappa \sum_{k=1}^{n_\text{s}} c_k\,\|\omega_k\|_0 \tag{25}$$

If n_s is large, problem (25) is very difficult to solve due to combinatorial explosion. Therefore, we perform a relaxation on the domain of definition of $\boldsymbol{\omega}$ and replace the discontinuous penalty by a smooth function P_δ, where $\delta \in (0,1]$, which converges to the ℓ_0-"norm" for $\delta \to 0$. For $\delta = 1$ this function has the form $P_{\delta=1}(\boldsymbol{\omega},\boldsymbol{c}) \coloneqq \boldsymbol{c}^{\top}\boldsymbol{\omega}$, otherwise

$$P_\delta(\boldsymbol{\omega},\boldsymbol{c}) \coloneqq \sum_{k=1}^{n_\text{s}} c_k\,f_\delta(\omega_k), \qquad \text{for } \delta \in (0,\tfrac{1}{2}), \tag{26}$$

where $f_\delta(x) : [0,1] \mapsto [0,1]$ is continuously differentiable and approximates the ℓ_0-"norm", see [1] for more details. The relaxed optimization problem

$$\min_{\boldsymbol{\omega}\in[0,1]^{n_\text{s}}} \quad \Psi\left[\,\boldsymbol{C}_{\text{post}}\left(\bar{\boldsymbol{\theta}}_0,\boldsymbol{\omega}\right)\right] + \kappa\,P_\delta(\boldsymbol{\omega},\boldsymbol{c}) \tag{27}$$

is first solved for $\delta = 1$ and then by a reiteration scheme for diminishing δ the optimal sensor weights $\boldsymbol{\omega}_{\text{opt}}$ tend to become sparse and $\{0,1\}$-valued for a suitable choice of $\kappa > 0$, cf. [1, 2]. We solve problem (27) by standard BFGS-SQP methods [20].

According to [12], the most prominent design criteria Ψ measuring the size of a matrix $\boldsymbol{C}$ are the following:

$$\Psi_\text{A} = \text{trace}(\boldsymbol{C}), \qquad \Psi_\text{D} = \det(\boldsymbol{C}), \qquad \Psi_\text{E} = \lambda_{\max}(\boldsymbol{C}). \tag{28}$$

It is known that problem (27) with $\delta = 1$ is convex for $\Psi = \Psi_\text{A}$ and $\Psi = \Psi_\text{D}$, see [17]. However, using Ψ_E requires non-smooth methods [13]. In this paper, we choose to compute the trace of the posterior covariance matrix.

4 NUMERICAL RESULTS FOR A REAL-SIZED HEATING NETWORK

We demonstrate our method for the heating grid in the district Darmstadt-Nord of the German city Darmstadt. The network is operated by the ENTEGA AG, the heat is provided by a central heating plant and distributed to 67 consumers through a pipe network of approximately 14.2 km length. The graph representation of the grid consists of 372 edges and 306 nodes.

Over the course of a day significant changes of overall demands as well as the relative distribution between the consumers is observed in practice. In the following we analyze the experiment design for each hour of the day individually and compare the results. Final investment decisions would have to be taken such that they yield good performance in all hours of the day. The hourly index is omitted in this section.

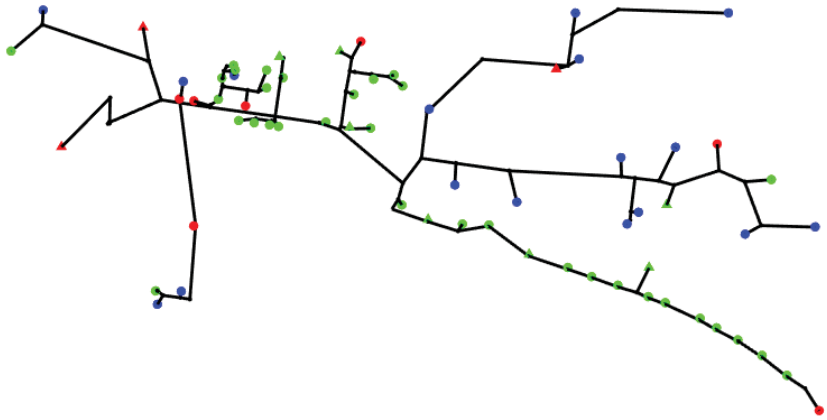


Figure 1: Network structure of the considered heating grid. Consumers with hourly measured demands are marked as circles, unmeasured as triangles. Green markers represent residential buildings, red ones commercial buildings and blue represent buildings in the group others.

4.1 Prior demand estimation

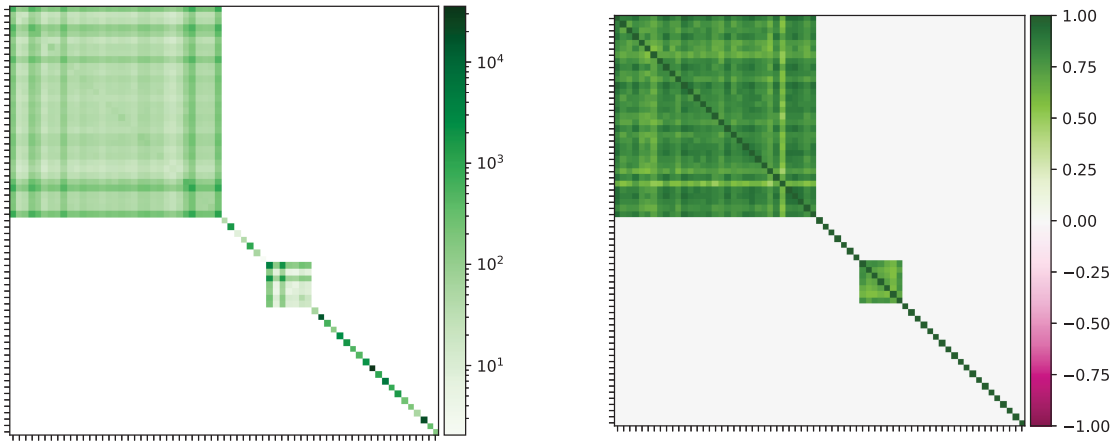
The prior load distribution $\mathcal{N}(\boldsymbol{\theta}_0, \boldsymbol{\Gamma}_0/\gamma)$ is estimated based on available consumption data. For 56 consumers the load is directly measured with hourly resolution. Let $\boldsymbol{K}_\mathrm{M}$ be the index set of all measured consumers and let $\boldsymbol{K}_\mathrm{U}$ be the index set of all unmeasured. The mean parameter $\theta_{0,i}$ for the measured demand $\dot{Q}_i^d$ of consumer i is given by the mean consumption over all observed days d

$$\theta_{0,i} = \frac{1}{D} \sum_{d=1}^{D} \dot{Q}_i^d, \qquad \text{for all } i \in \boldsymbol{K}_\mathrm{M}\,. \tag{29}$$

Unmeasured consumers $j \in \boldsymbol{K}_\mathrm{U}$ are paired with similar measured consumers $j^* \in \boldsymbol{K}_\mathrm{M}$. The expected demand is given by scaling the expectation value of the paired consumer by the associated total consumption that was also available for us, $\dot{Q}_j^{2019}$, respectively, $\dot{Q}_{j^*}^{2019}$:

$$\theta_{0,j} = \frac{\dot{Q}_j^{2019}}{\dot{Q}_{j^*}^{2019}}\, \theta_{0,j^*}, \qquad \text{for all } j \in \boldsymbol{K}_\mathrm{U}\,. \tag{30}$$

The return temperatures $\boldsymbol{T}^\mathrm{set}$ for each consumer are estimated by the same procedure. Different heat consumption profiles can be explained by the varying outdoor temperatures and by different consumer behavior [6]. Grouping consumers for whom similar behavior is expected leads to three groups. The first group $\boldsymbol{G}^\mathrm{res}$, being residential buildings, consist of 40 multi-family houses and one single-family house. The second group $\boldsymbol{G}^\mathrm{TC}$ is made up by 10 commercial buildings. The remaining 17 buildings form the group $\boldsymbol{G}^\mathrm{other}$ and are not expected to show strong similarities. Fig. 1 shows the distribution of demand classes in the network.



(a) covariance matrix (b) correlation matrix

Figure 2: In 2a the covariance matrix $\boldsymbol{\Gamma}_0$ for the heat demand $\dot{Q}$ hour 3 to 4 pm is shown. The first block represents the group of measured residential buildings, followed by the unmeasured residential buildings for which all correlation coefficients are set to zero. The second block represents measured and unmeasured commercial buildings, the third the buildings of the "other" group. The variances can differ by several orders of magnitude due to the largely different size of the buildings. 2b shows the corresponding correlation matrix normalized by the standard deviations.

The standard deviation σ_i for the measured demands i are estimated by

$$\sigma_i = \sqrt{\frac{1}{D}\sum_{d=1}^{D}\left(\dot{Q}_i^d - \theta_{0,i}\right)^2}, \qquad \text{for all } i \in \boldsymbol{K}_\mathrm{M}\,. \tag{31}$$

Within the first two groups the normalized standard deviations $\sigma_i/\theta_{0,i}$ are close-by for all hourly measured consumers. The standard deviation of the unmeasured demands

$$\sigma_j = \frac{\theta_{0,j}}{|\boldsymbol{K}_\mathrm{M}|}\sum_{i\in\boldsymbol{K}_\mathrm{M}}\frac{\sigma_i}{\theta_{0,i}}, \qquad \text{for all } j \in \boldsymbol{K}_\mathrm{U}, \tag{32}$$

are therefore estimated by scaling the mean normalized variations by the estimated demand expectations. For the group $\boldsymbol{G}^\mathrm{other}$ all standard deviations can be gathered directly from the data. The entries $\Gamma_{0,ij}$ for the prior covariance matrix $\boldsymbol{\Gamma}_0$ can now be gathered by

$$\Gamma_{0,ij} = \begin{cases} \sigma_i^2, & \text{if } i = j, \\ \frac{1}{D-1}\sum_{d=1}^{D}\left(\dot{Q}_i^d - \theta_{0,i}\right)\left(\dot{Q}_j^d - \theta_{0,j}\right), & \text{if } i,j \in \boldsymbol{G}^\mathrm{res}\cap\boldsymbol{K}_\mathrm{M} \text{ or } i,j \in \boldsymbol{G}^\mathrm{TC}\cap\boldsymbol{K}_\mathrm{M}, \\ 0, & \text{else.} \end{cases} \tag{33}$$

The covariance coefficients between two groups as well as for consumers in group $\boldsymbol{G}^\mathrm{other}$ are set to 0, because no similarity in the demand patterns are expected here. When ordered according to the groups, the block structure of the covariance matrix becomes clearly visible as it is shown in Fig. 2a for the hour 3 to 4 pm.

The factor $\gamma > 0$ must be large enough to achieve regularity of the covariance matrix, see eq. (24), and to make the solver of the optimization problem (27) numerically stable. This factor also brings the prior into proper scaling when added to the data misfit part of the covariance formula in eq. (24). In our case, taking $\gamma = 5 \times 10^3$ was sufficient.

4.2 Sensors

Two different kinds of sensors are considered, namely pressure sensors and power flow sensors measuring three values simultaneously. There are two parallel pipes in the grid, one delivering hot water and one returning the cold water to the heating plant. Let $(i, j) \in \boldsymbol{E}^{\text{pipe}}$ be a pipe on the return side (cold water) and $(k, l) \in \boldsymbol{E}^{\text{pipe}}$ be the corresponding parallel pipe on supply side (hot water). A potential power flow sensor would then measure the mass flow $\dot{m}_{ij}$, which is equal for both pipes, as well as the temperatures T_{ij}^{end} and T_{kl}^{start}. The power transmitted along the pipe-pair is given by the temperature difference between the two pipes and the mass flow through them:

$$P_{ij}^{kl} = c_p \dot{m}_{ij} \left(T_{ij}^{\text{end}} - T_{kl}^{\text{start}} \right). \tag{34}$$

There are 150 plausible positions for such heating power sensors and as many pressure sensor locations as nodes in the network are available, altogether we obtain $n_{\text{s}} = 456$ candidate sensor positions.

Each sensor has a different accuracy when measuring the quantities of interest. The pressure sensors operate with a 0.04 % precision of the measured pressure value: $\Delta p_i = 0.04\,\% \cdot p_i$, for all $i \in \boldsymbol{V}$. The mass flow $\dot{m}_{ij}$ is determined by measuring the flow speed with a fixed accuracy. The accuracy of the mass flow therefore depends on the pipe diameter d_{ij} and the density of water ρ:

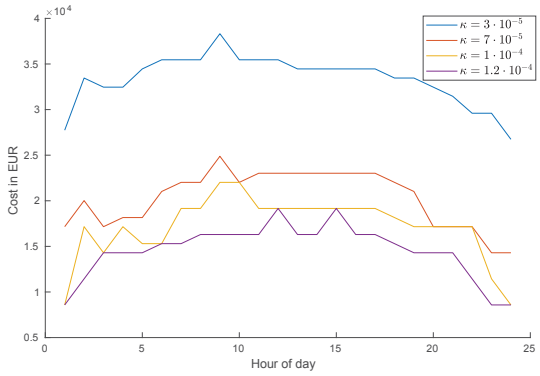
$$\Delta \dot{m}_{ij} = \rho \frac{d_{ij}^2}{4} \cdot 0.012\,\mathrm{m\,s^{-1}}, \qquad \text{for all } (i, j) \in \boldsymbol{E}\,. \tag{35}$$

For the temperature measurement, we have $\Delta T_i = \Delta T_{kl}^{\text{end}} = \Delta T_{kl}^{\text{start}} = 0.6\,^{\circ}\mathrm{C}$, for all $i \in \boldsymbol{V}$ and $(k, l) \in \boldsymbol{E}$. These values form the diagonal entries of the covariance matrix $\boldsymbol{\Sigma}$ of the noise model $\boldsymbol{\varepsilon}$, see eq. (17). The non-diagonal entries in $\boldsymbol{\Sigma}$ are set to zero since we assume the sensor readings to be statistically independent.

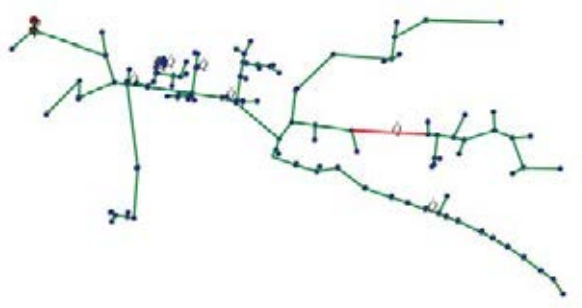
4.3 Computational results

We solved the state equation (15) by a Newton-method with projected gradients after the starting point had been computed by a truncated fixed-point iteration. The usage of projected gradients enhanced the convergence properties, since the network was designed to yield only positive mass flows $\dot{\boldsymbol{m}}$. After computing the sensitivity matrix $\boldsymbol{J}$, see eq. (21), we solved problem (27) with a standard SQP-method where the Hessian is constructed by BFGS-updates. The reiteration scheme with the penalty term was performed starting with $\delta_1 = 1$ and then updating $\delta_{k+1} := \delta_k / 2$ for $k = 1, \ldots, 6$. Thus, in a few reiterations, the optimal sensor weights became sparse and almost $\{0, 1\}$-valued. We additionally want to point out that the solution was found after approximately 120 function and 110 gradient evaluations in total for each hour, respectively.

We pick the hour of the day with the highest heat demand, which was 9 to 10 am, and compare the results for different values of κ in Tab. 1. We also generated 500 random vectors



(a) Costs of optimal sensors for different κ and hours of the day



(b) Optimized sensor positions for 7 am, $\kappa = 7 \times 10^{-5}$ and 9 am, $\kappa = 1 \times 10^{-4}$

Figure 3: 3a The optimized set of sensors depends on the chosen value for κ as well as the hour of the day due to different prior assumptions for the demands. Sets resulting in equal costs are usually very similar or equal as in 3b. This indicates that the proposed approach leads to sensor positions that perform well for all hours.

ω_{rnd} and computed the average of their respective Ψ_A, Ψ_D and Ψ_E criteria, to compare with an optimal selection of ω and with no sensors at all. We observe that the smaller the value of κ the more sensors are used and the smaller is the design criterion. However, the greater the number of installed sensors the higher the costs. The improvement of prior information about the demands which is achieved by using sensors at the optimized positions ω_{opt}, is between 56 % and 64 % when evaluating Ψ_A, clearly over 99 % when using Ψ_D and between 71 % and 77 % when evaluating Ψ_E. A random selection of an equal number of sensors produces a much smaller difference while the computational efforts are much higher.

As seen in Fig. 3a, the optimal trade-off between the cost of sensors and the design criterion varies for the different hours of the day for a constant κ resulting in a different number of sensors and therefore different costs at the optimal solution. This behavior can be attributed to the different prior assumptions over the demands for each hour of the day. Fig. 4 shows the design criteria Ψ_A over the cost for the sensors for optimized sensor networks for each hour individually. By changing the value for κ the optimum can be altered along these estimated optimality curves. It can be seen from this representation that the curve steeply decreases for low numbers of sensors and flattens for higher numbers. This is similar to a pareto-front known from multi-objective optimization. Additionally, it seems as if the optimality curves tend to

Table 1: Comparison of the solution of problem (27) for different κ from hour 9 to 10 am. We additionally computed the design criteria Ψ_D and Ψ_E.

#	$\|\omega\|_0$	$c^\top \omega$	κ	Ψ_A	Ψ_D	Ψ_E
$\emptyset$	0	-	-	23.46	1.65×10^{-127}	7.91
ω_{rnd}	7	-	-	22.60	1.20×10^{-127}	7.62
ω_{opt}	7	1.630×10^4	1.2×10^{-4}	10.27	3.48×10^{-131}	2.28
ω_{rnd}	9	-	-	22.58	1.14×10^{-127}	7.65
ω_{opt}	9	2.202×10^4	0.7×10^{-4}	9.28	3.45×10^{-132}	2.00
ω_{rnd}	15	-	-	21.97	8.98×10^{-128}	7.47
ω_{opt}	15	3.546×10^4	0.3×10^{-4}	8.55	5.55×10^{-133}	1.80

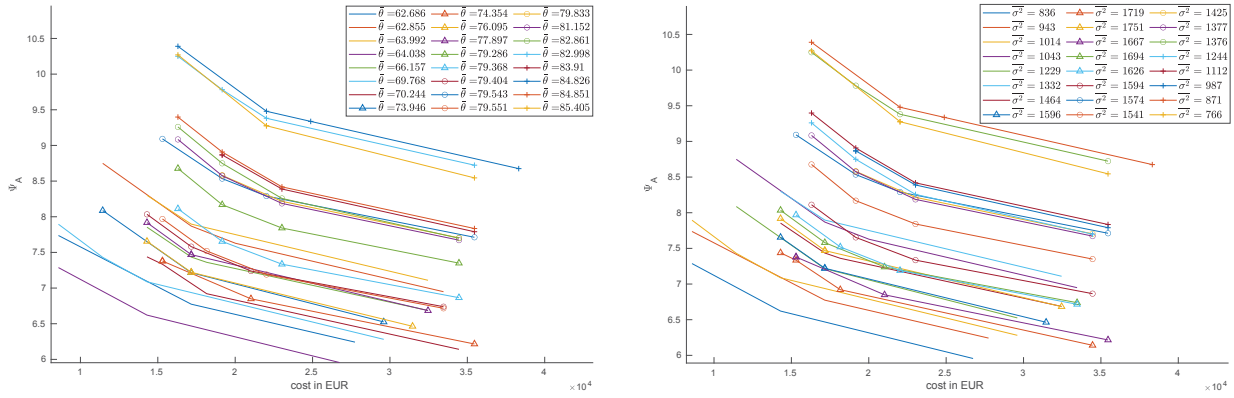


Figure 4: By choosing κ, a trade-off can be made between a better information score Ψ_A and lower costs. The lines along the weighting moves tend to shift towards higher prices and worse scores for increasing $\boldsymbol{\theta}_0$ and σ^2. The values given in these figures are the mean values for all consumers. It can be seen that there are also some exceptions to this trend.

shift towards higher costs as well as higher Ψ_A-values with overall higher demands and higher variances for the demands. Even though these trends do not hold true for all demand priors examined, it should be taken into account when choosing a κ for practical applications. Fig. 3b shows the position of sensors for the hour 7 am and $\kappa = 7 \times 10^{-5}$ which is exactly the same as the set for the hour 9 am choosing $\kappa = 1 \times 10^{-4}$. Similarly as in this example the chosen sensor positions for different hours tend to match each other closely if κ is set in such a way that similar cost arise in the optimum. It can therefore be expected that a set of sensors optimized for one hour provides good results for the other hours as well.

The mapping $\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{\theta})$ can be used to estimate the network's state $\boldsymbol{y}$. Since we assume normal distributed uncertainty for the consumers demand, we can use the linearised model to estimate an normal distribution for the state $\boldsymbol{y}_{\text{est}} \in \mathcal{N}\left(\boldsymbol{y}_0, \boldsymbol{\Sigma}_{\boldsymbol{y}}\right)$ with the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{y}}$ given by

$$\boldsymbol{\Sigma}_{\boldsymbol{y}} \propto \boldsymbol{J}\, \boldsymbol{C}_{\text{post}}\, \boldsymbol{J}^{\top}, \tag{36}$$

where $\boldsymbol{J}$ is the sensitivity matrix from (21). In Fig. 5 the estimated states for the prior and the posterior estimation over the demands for hour 9 am and $\kappa = 1 \times 10^{-4}$ are compared, by calculating the change in variance for each individual state variable. The graphic shows the temperatures variance as node color and the mass flow variance as edge color for the supply side of the network. It can be seen directly, that the uncertainty does not only decrease near the measurements, but for almost all state variables. When separating the grid into different sections it seems as if within each section the gains are relatively even, but differ strongly between the sections. This can be motivated by the different consumer structure. By comparing with Fig. 1 it can be observed that the lowest section on the right side mostly consists of residential buildings. For these the variances were already rather low for the prior assumption as seen in Fig. 2a. In the middle section on the right side, many consumers are classified as "others", showing comparable large demands and large variations in the prior. Therefore the relative gain is larger for this area.

Figure 5: The linearized model can be used as a state estimator. Using the posterior estimation for the consumers demand significantly reduces the uncertainty compared to the prior assumption. This plot shows the uncertainty change for the temperatures (node color) and the mass flow (edge color) for the supply side of the network.

5 CONCLUSION

In this paper we presented a Bayesian approach for a sensor placement problem in heating networks to improve the reliability of our knowledge about one of the most important system parameters – the consumer's heat consumption. The sensor placement task can be interpreted as an optimal experimental design problem which we modeled using the solution of the state equation of the heating network, its sensitivity matrix and the Bayes formula for the parameter's covariance. We solved this optimization problem by a BFGS-SQP method with a reiteration scheme to obtain sparse and almost $\{0,1\}$-valued sensor weights. The optimally positioned sensors measure temperatures, pressures and mass flows in the network at a trade-off between small covariance values and low costs. In subsequent numerical experiments we applied our method to the heating network of the northern district of the German city Darmstadt. We showed that the optimal placement of a few sensors significantly increased the informational value of the uncertain demands at low cost when compared to randomly placed sensors. This shows that our method is superior to Monte-Carlo approaches. An enhanced knowledge of the demand values provides the basis for a detailed monitoring of the network state in order to reduce industrial waste-heat.

Even though the numerical results are very promising, some barriers remain for practical use of this approach. In order to optimize the sensor positions, a prior distribution for the demands is required. For our investigated network, this was done by analyzing each consumer's past demand which was measured with hourly resolution for most buildings. However, measurements of this kind are rather uncommon in district heating networks. The expected demand as well as the variance of the demands who were not measured, were estimated by comparing the buildings with measured ones. In order to treat these consumers equally to the measured ones, the non-

diagonal entries of the covariance matrix should be estimated as well. However, this task is not trivial since estimating the coefficients, e.g., by the mean of the coefficients in the same buildings group, will most likely lead to a matrix which is not positive definite and therefore no valid covariance matrix.

The optimal sensor positions found in our setting may differ from the sensors that would be chosen in real life application. For example, sensors are still placed in the connection lines between single buildings and the main grid, which is equivalent to directly measuring the demand at the consumers heat exchange station. These measurement positions would be unsuitable for real life applications due to fast demand changes that are likely to occur. However, this possibility is not considered by the steady state model. Another example is given by the two pressure sensors in Fig. 3b, which are next to each other. In our model, this is equivalent to measuring the mass flow through this pipe as given by (9). For practical applications this would not be a preferable setup, as the pressure differences between the two nodes are too small to gain usable information.

In order to install sensors in real district heating grids, many additional factors need to be taken into account, e.g., how well a potential sensor position could be reached, for installation or maintenance. The optimization scheme proposed in this paper can contribute to this decision process by suggesting optimal sensor positions under simplified conditions or by comparing different possible settings in the context of the chosen model.

ACKNOWLEDGMENTS

We would like to thank the German Research Foundation (DFG) – project number 57157498 – CRC 805 and the German Federal Ministry for economic affairs and energy (BMWi) – project number 03EN3012A – for funding this research. We kindly acknowledge the help of ENTEGA AG who provided the necessary data and helped interpreting the results.

REFERENCES

[1] Alen Alexanderian, Noemi Petra, Georg Stadler, and Omar Ghattas. A-optimal design of experiments for infinite-dimensional Bayesian linear inverse problems with regularized ℓ_0-sparsification. *SIAM Journal on Scientific Computing*, 36(5):A2122–A2148, 2014.

[2] Alen Alexanderian, Noemi Petra, Georg Stadler, and Omar Ghattas. A fast and scalable method for A-optimal design of experiments for infinite-dimensional Bayesian nonlinear inverse problems. *SIAM Journal on Scientific Computing*, 38(1):A243–A272, 2016.

[3] Irene Bauer, Hans G. Bock, Stefan Körkel, and Johannes P. Schlöder. Numerical methods for optimum experimental design in dae systems. *Journal of Computational and Applied Mathematics*, 120(1-2):1–25, 2000.

[4] Ilyes Ben Hassine and Ursula Eicker. Impact of load structure variation and solar thermal energy integration on an existing district heating network. *Applied Thermal Engineering*, 50(2):1437–1446, 2013.

[5] Hanmin Cai, Shi You, and Jianzhong Wu. Agent-based distributed demand response in district heating systems. *Applied Energy*, 262:114403, 2020.

[6] Bronislav Chramcov. Heat demand forecasting for concrete district heating system. *International Journal of Mathematical Models and Methods in Applied Sciences*, 2010.

[7] Jean Duquette, Andrew Rowe, and Peter Wild. Thermal performance of a steady state physical pipe model for simulating district heating grids with variable flow. *Applied Energy*, 178:383–393, 2016.

[8] ENTEGA AG, Darmstadt, Germany. *Technische Anschlussbedingungen (TAB) Fernwärme Darmstadt*, 2017.

[9] Tingting Fang and Risto Lahdelma. State estimation of district heating network based on customer measurements. *Applied Thermal Engineering*, 73(1):1211–1221, 2014.

[10] Valerii V. Fedorov and Sergei L. Leonov. *Optimal Design for Nonlinear Response Models*. CRC Press, 1st edition, 2013.

[11] Eric B. Flynn and Michael D. Todd. A Bayesian approach to optimal sensor placement for structural health monitoring with application to active sensing. *Mechanical Systems and Signal Processing*, 24(4):891–903, 2010.

[12] Gaia Franceschini and Sandro Macchietto. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63(19):4846–4872, 2008.

[13] Jean-Baptiste Hiriart-Urruty and Adrian S. Lewis. The Clarke and Michel-Penot subdifferentials of the eigenvalues of a symmetric matrix. *Computational Optimization and Applications*, 13(1):13–23, 1999.

[14] Stefan Körkel, Ekaterina Kostina, Hans G. Bock, and Johannes P. Schlöder. Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes. *Optimization Methods and Software*, 19(3-4):327–338, 2004.

[15] Xuezhi Liu, Jianzhong Wu, Nick Jenkins, and Audrius Bagdanavicius. Combined analysis of electricity and heat networks. *Applied Energy*, 162:1238–1250, 2016.

[16] Henrik Lund, Sven Werner, Robin Wiltshire, Svend Svendsen, Jan Eric Thorsen, Frede Hvelplund, and Brian Vad Mathiesen. 4th generation district heating (4gdh): Integrating smart thermal grids into future sustainable energy systems. *Energy*, 68:1–11, 2014.

[17] Ira Neitzel, Konstantin Pieper, Boris Vexler, and Daniel Walter. A sparse control approach to optimal sensor placement in PDE-constrained parameter estimation problems. *Numerische Mathematik*, 143(4):943–984, 2019.

[18] Tinsley Oden, Robert Moser, and Omar Ghattas. Computer predictions with quantified uncertainty, Part I. *SIAM News*, 43(9):1–3, 2010.

[19] Michael Papadopoulos and Ephrahim Garcia. Sensor placement methodologies for dynamic testing. *AIAA journal*, 36(2):256–263, 1998.

[20] Michael J. D. Powell. A fast algorithm for nonlinearly constrained optimization calculations. *Lecture Notes in Mathematics*, 630:144–157, 1978.

[21] Friedrich Pukelsheim. *Optimal Design of Experiments*. Society for Industrial and Applied Mathematics, 2006.

[22] Andrew M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 19:451–559, 2010.

[23] Guoqiang Sun, Wenxue Wang, Yi Wu, Wei Hu, Zijun Yang, Zhinong Wei, Haixiang Zang, and Sheng Chen. A nonlinear analytical algorithm for predicting the probabilistic mass flow of a radial district heating network. *Energies*, 12(7):1215, 2019.

[24] Shoujun Zhou, Zheng O'Neill, and Charles O'Neill. A review of leakage detection methods for district heating networks. *Applied Thermal Engineering*, 137:567–574, 2018.