Sensor Placement and State Estimation in Water Distribution Systems Using Edge Gaussian Processes

Kerimov, Bulat; Pons, Vincent; Pritsis, Spyros; Taormina, Riccardo; Tscheikner-Gratl, Franz

DOI
10.3390/engproc2024069150

Publication date
2024

Document Version
Final published version

Published in
Engineering Proceedings

engineering proceedings

MDPI

Proceeding Paper

Sensor Placement and State Estimation in Water Distribution Systems Using Edge Gaussian Processes †

Bulat Kerimov [1,*], Vincent Pons [1], Spyros Pritsis [1], Riccardo Taormina [2] and Franz Tscheikner-Gratl [1]

1 Department of Civil and Environmental Engineering, Norwegian University of Science and Technology-NTNU, Torgarden, 7031 Trondheim, Trøndelag, Norway; vincent.pons@ntnu.no (V.P.); spyridon.pritsis@ntnu.no (S.P.); franz.tscheikner-gratl@ntnu.no (F.T.-G.)
2 Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, 2600 GA Delft, The Netherlands; r.taormina@tudelft.nl
* Correspondence: bulat.kerimov@ntnu.no
† Presented at the 3rd International Joint Conference on Water Distribution Systems Analysis & Computing and Control for the Water Industry (WDSA/CCWI 2024), Ferrara, Italy, 1–4 July 2024.

Abstract: The operation of water distribution systems is based on reliable knowledge about the steady state of the system. This involves sensors to measure flow, facilitating a comprehensive overview of the system's performance. Given the costs associated with sensor installation and operation, it is important to be strategic with sensor allocation. Recently developed Gaussian Processes with topological kernels can efficiently model mass and energy conservative flows and provide uncertainty bounds. Our work proposes a novel method of state estimation and a greedy search algorithm for water flow meter placement based on the uncertainty bounds provided by a Gaussian Process.

Keywords: sensor placement; flowrate estimation; Gaussian processes; uncertainty

Citation: Kerimov, B.; Pons, V.; Pritsis, S.; Taormina, R.; Tscheikner-Gratl, F. Sensor Placement and State Estimation in Water Distribution Systems Using Edge Gaussian Processes. *Eng. Proc.* **2024**, *69*, 150. https://doi.org/10.3390/engproc2024069150

Academic Editors: Stefano Alvisi, Marco Franchini, Valentina Marsili and Filippo Mazzoni

Published: 19 September 2024

1. Introduction

The main challenge of planning a monitoring scheme for a water distribution system (WDS) is striking a balance between maximizing the observability and the cost. Most attempts to tackle this challenge use computationally expensive optimization processes that can be prohibitive for large networks [1]. Approaches based on graph theory provide a remedy; however, they do not account for physical laws that govern the system [2].

Recent advancements in geometric deep learning enable the incorporation of network topology and physical laws into the model [3]. This includes novel Gaussian Processes with topological kernels that can efficiently model mass and energy conservative flows. By including connectivity information into the kernel, Gaussian Processes (GPs) can interpolate flows in a network based on mass-conservation principles. Consequently, the variance of the GP can serve as an indicator of uncertainty. Our work proposes a greedy search algorithm for water flow meters based on the uncertainty of the estimated flowrates.

2. Materials and Methods

2.1. Problem Statement

We define a directed graph $G = (V, E)$ that describes a water distribution system. The set of nodes V corresponds to the junction, reservoirs, and tanks, while the set of edges E represents the pipes. The direction of the edges is chosen arbitrarily. The matrix $B_1 \in R^{|V|\times|E|}$ denotes the incidence between edges and nodes of the graph [4]. We additionally define a set of loops C in the network and assign an arbitrary orientation. Similarly, $B_2 \in R^{|E|\times|C|}$ is an incidence between the cells and incidental edges.

We assume noisy demands on the nodes $q \in R^{|N_c|}$ and flowrate sensors, where N_c is the number of consumption nodes, and a set of k sensor edges, i.e., water flow meters. The desired output is the vector of flowrates on the pipes $f \in R^{|E|-k}$ along with the uncertainty bounds. The negative sign of the vector value means that the true direction of the flow is the opposite to the one selected. We are additionally interested in finding the set S of k best sensor locations out of set E.

2.2. Flowrate Estimation with Gaussian Processes

Let X be a set of pipes. A flowrate vector can be modeled with a multi-dimensional Gaussian process $f \sim GP(\mu, k)$, with mean vector $\mu = \mu(x)$ and covariance matrix $K = R^{|E|\times|E|}$. We assume that the prior mean μ is zero. The resulting function on the edges $f(x)$ is assumed to follow a multivariate Gaussian distribution where each dimension in it corresponds to a pipe.

For a given set of sensor locations and sensor measurements (x_i, y_i), we define the model $y_i = f(x_i) + \epsilon_i$, where $f \sim GP(0, k)$ and noise $\epsilon_i \sim N(0, \sigma^2)$. The posterior of f, given the observations at x_i, is another GP. Its conditional mean and covariance at the rest of the pipes (.) are as follows:

$$\mu|_y(.) = K_{.x}\left(K_{xx} + \sigma^2 I\right)^{-1} y \quad (1)$$

$$K(.,.\prime)|_y = k(.,.\prime) - K_{.x}\left(K_{xx} + \sigma^2 I\right)^{-1} K_{x.\prime} \quad (2)$$

The result of the posterior is a multi-variate distribution of flowrates, conditioned by known flowrates at sensor locations. The predicted flowrates can be extracted as the mean of the Gaussian flowrates.

2.2.1. Mass-Conservative Covariance

The mass conservation principle states that all incoming water flows are equal to the outflow.

$$B_1 f = q \quad (3)$$

Recent works introduce a kernel that describes the relationship between the values defined on the edges of the graph. In fact, if the covariance is selected as $L^H_{down} = e^{-B_1^T B_1}$, the sampled f will be mass conservative [3].

To ensure energy conservation, the sampled flowrates are transformed using the Hazen–Williams relation into headlosses h. As the relation is non-linear, the distribution of h is no longer Gaussian. Energy conservation can be assured by projecting h with a linear operator $L^H_{up} = e^{B_2 B_2^T}$ [5].

$$h' = L^H_{up} h \quad (4)$$

The resulting h' are energy conservative, meaning that the sum of energy losses within a closed loop with respect to their orientations are equal to 0.

2.2.2. Virtual Edges

To incorporate the demand on the nodes into the estimation, we propose to extend the graph G with $|V|$ virtual nodes paired with each real node. As a result, q on the corresponding nodes act as sensory input on pipes to GP.

2.3. Sensor Placement

The sensor placement strategy is based on the greedy selection of the GP dimension with the highest variance. The algorithm iteratively updates the set of sensors S and evaluates $\mu|_y$ and $K|_y$, applying Equations (1) and (2). The latter estimates the uncertainty of the prediction. This strategy is akin to the Bayesian optimization procedure [6].

2.4. Evaluation

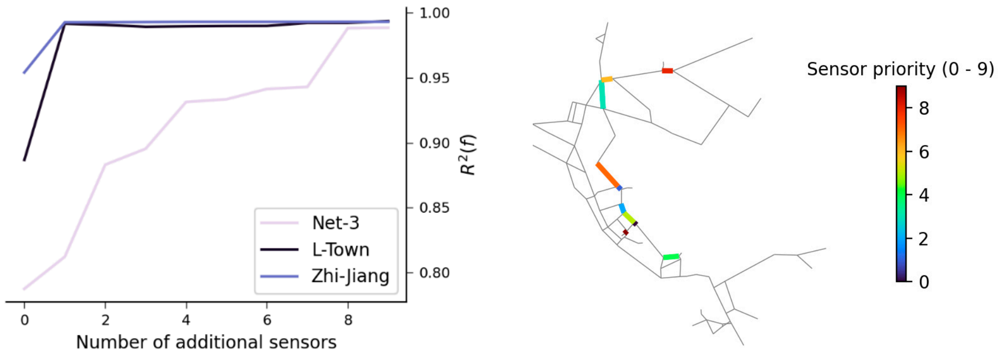
Algorithm 1 is evaluated based on 3 benchmark networks: Zhi-Jiang [7], area C of L-Town [8], and Net-3 [9]. The initial set S of sensors is selected as the flowrates on virtual edges (i.e., demands and reservoir outflows), $k = 10$. For Net-3, we additionally visualize the location of the sensors. The accuracy of the state estimation is evaluated with the R^2 of flowrates on unknown pipes against the ones obtained via EPANET [9].

Algorithm 1. Sensor Placement Strategy

Input: Covariance $\mathbf{K}$, k, $x \subset E$, S
for j = 1 **to** k:
Update $\mathbf{K}$ according (1) and (2)
$h(x) \leftarrow HazenWilliams(f(x))$
$h\prime(x) \leftarrow L_{up}^{H} h(x)$
$f(x) \leftarrow HazenWilliams^{-1}(h(x))$
$S_x^{argmax} = Var(f(x))$
$S \leftarrow S \cup S'$
end

3. Results

Figure 1a visualizes the dynamics of errors of the estimated flowrates against the number of flow meters installed in the network. The estimation accuracy increases with the number of sensors. For ZJ and L-Town, the largest gain is visible with the first suggestion as the accuracy reaches 0.99 in R^2. However, further suggestions are less informative. In the example of Net-3, the accuracy gradually increases with the number of flowrate sensors and reaches 0.97.



(**a**) (**b**)

Figure 1. Results of a sensor placement algorithm. (**a**) Ten selected sensor locations on the example of Net-3. (**b**) Accuracy of the flowrate estimation.

4. Discussion

For ZJ and L-Town, the sensor placement algorithm prioritizes the allocation of sensors on the pipes that are located next to the reservoir. Assuming the demands are known with some noise factor, these pipes are the most influential for estimating water flow in the pipes. Figure 1b shows that the algorithm prioritizes the pipes that act as bottlenecks and influence the rest of the network.

The difference in the accuracy between the networks requires further investigation.

Currently, the algorithm shows the preliminary results on simple networks and requires further investigation. Further work can include pressure sensor placement and can combine multiple Gaussians to incorporate energy conservation more thoroughly.

5. Conclusions

This work introduces a novel method of estimation of flowrates and uncertainty bounds based on GP. We additionally propose a greedy water flow sensor placement algorithm that systematically prioritizes pipes in the water network with the highest uncertainty of estimated flowrates. This approach is similar to the exploration objective based on the acquisition function of Bayesian optimization. For simple networks, this allows for interpolating the flowrates with assumed demands and a single flowrate sensor with high accuracy. Finally, by omitting the need to run expensive hydraulic simulations, we provide a fast sensor-placement technique that is both geometrically and physically informed.

Author Contributions: Conceptualization, B.K., F.T.-G. and R.T.; methodology, B.K., S.P., V.P. and F.T.-G., software, B.K.; validation, B.K. and F.T.-G.; formal analysis, B.K.; investigation, B.K.; resources, B.K.; data curation, B.K.; writing—original draft preparation, B.K. and S.P.; writing—review and editing, B.K., S.P. and F.T.-G.; visualization, B.K.; supervision, F.T.-G. and R.T.; project administration, F.T.-G. and R.T.; funding acquisition, F.T.-G. and R.T. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the financial support through the NTNU Green2050 Centre for Green Shift in the Built Environment and the TU Delft AI Labs program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kang, D.; Lansey, K. Optimal Meter Placement for Water Distribution System State Estimation. *J. Water Resour. Plan. Manag.* **2010**, *136*, 337–347. [CrossRef]
2. Perelman, L.; Ostfeld, A. Application of Graph Theory to Sensor Placement in Water Distribution Systems. In Proceedings of the World Environmental and Water Resources Congress 2013, Cincinnati, OH, USA, 19–23 May 2013.
3. Yang, M.; Borovitskiy, V.; Isufi, E. Hodge-Compositional Edge Gaussian Processes. *arXiv* **2023**, arXiv:2310.19450.
4. Todini, E.; Rossman, L.A. Unified Framework for Deriving Simultaneous Equation Algorithms for Water Distribution Networks. *J. Hydraul. Eng.* **2013**, *139*, 511–526. [CrossRef]
5. Alain, M.; Takao, S.; Paige, B.; Deisenroth, M.P. Gaussian Processes on Cellular Complexes. *arXiv* **2023**, arXiv:2311.01198.
6. Nguyen, V. Bayesian optimization for accelerating hyper-parameter tuning. In Proceedings of the IEEE 2nd International Conference on Artificial Intelligence and Knowledge Engineering (AIKE 2019), Sardinia, Italy, 3–5 June 2019. [CrossRef]
7. Dandy, G. 06 Zhi Jiang. *Int. Syst.* **2016**, *6*. Available online: https://uknowledge.uky.edu/wdst_international/ (accessed on 7 September 2024).
8. Vrachimis, S.G.; Eliades, D.G.; Taormina, R.; Kapelan, Z.; Ostfeld, A.; Liu, S.; Kyriakou, M.; Pavlou, P.; Qiu, M.; Polycarpou, M.M. Battle of the leakage detection and isolation methods. *J. Water Resour. Plan. Manag.* **2022**, *148*, 04022068. [CrossRef]
9. Rossman, L.A. 06 EPANET Net 3. Software Manual Examples. 2016. Available online: https://uknowledge.uky.edu/wdst_manuals/2 (accessed on 7 September 2024).