Vol. 11, Special Issue (Theveli 2023), pp. 107-116

RESEARCH REPORTS

Improving Water Distribution Network Models with Data Assimilation: Unleashing the Power of Real-Time Improvements

IBRAHIM MIFLAL FAYAZ, MARIO CASTRO-GAMA, & LEONARDO ALFONSO SEGURA

IHE Delft Institute for Water Education, The Netherlands

ABSTRACT *Ensuring a reliable water supply in the face of changing conditions and growing demand is a critical global challenge. Water distribution networks (WDNs) are essential infrastructure, but traditional modeling methods based on historical data often struggle to adapt in real-time and integrate new information. In order to lower model errors in WDN models, this study explores the use of a Data Assimilation (DA) method that makes use of the Ensemble Kalman Filter. The study explores the effectiveness of the DA method, and a novel Greedy Algorithm (GA) for optimizing the location of sensors. The study shows that the DA method improves the WDN model and the GA is able to successfully determine optimal sensor locations. It was observed that increasing the number of sensors in the WDN increased the effectiveness of the DA method. This study highlights the potential of data assimilation to improve WDN modeling. Water utilities can gain from more precise predictions, greater system performance, and efficient decision-making for the management and maintenance of water distribution networks by allowing models to adapt and improve dynamically.*

Keywords: Water Distribution Networks, Hydraulic Modelling, Ensemble Kalman Filter, Optimization, Greedy Algorithm

Introduction

Water is an essential element for life, and its management and accessibility are pressing global concerns. According to the United Nations (2022a), the availability of water varies worldwide, with some regions facing scarcity and others experiencing surpluses, compounded by the growing demand for water from expanding populations and economic growth. The 6th United Nations (UN) Sustainable Development Goal (SDG) aims to provide access to clean water and sanitation for all by 2030. Its fourth target prioritizes improving water efficiency and sustainable management of fresh water resources to combat water scarcity (United Nations, 2022a).

A WDN is a collection of interconnected pipes, valves, pumps, tanks, and other equipment that carries water from a source, such as a treatment plant, to end users, like residences and commercial buildings. A WDN ensures that all users have constant access to pressurized, safe drinking water. WDNs play a crucial role in delivering potable water to consumers and are vital to the community and urban infrastructure. To maximize the performance of WDNs, numerous methods and tools have been developed, including hydraulic models, demand management

ISSN 2308-5959/20231130

tactics, and real-time control and monitoring systems. These methods can be used to improve the effectiveness, efficiency, and sustainability of WDNs.

Hydraulic modelling is a commonly used decision support tool by water operators to simulate and analyze WDN behaviour. This involves creating a digital representation of the WDN using computer software, simulating various scenarios, and evaluating the impact on the system. The model considers factors such as water pressure and flow rate to predict the behaviour of the WDN. The results of simulations help optimize its design, operation, and maintenance to ensure reliability and efficiency. The limitations of offline models in capturing the intricate nature and diversity of real-world systems have been acknowledged in previous studies (Antonowicz et al., 2018; Rossman, 1994). Data Assimilation emerges as a viable solution to overcome these challenges.

Data Assimilation

A typical Data Assimilation (DA) process involves comparing a model forecast (prior) with recently acquired observations, updating the model state to reflect the observations (posterior), starting a new forecast, and so on ECMWF (2013). DA is important, as a model which is running independently of measured data is likely to diverge away from reality due to unknowns in the process and due to uncertainties in the initial conditions (University of Reading, 2022).

Measurements and models are combined through DA in a way that takes into consideration each other's uncertainties while also adhering to specific constraints. These consist of how the measurements physically connect to the system's variables and the rules of the system as expressed by the model equations (University of Reading, 2022).

When there is a high variance in the measured data, the assimilated model results are biased towards the model estimates which have a lower variance. Similarly, when there is a high variance in the model estimates, the assimilated model is biased towards the measured data which has a lower variance (Mathworks, 2022).

Digital Transformation

The ongoing global digital transformation is revolutionizing various industries, including water distribution. Installation of Advanced Metering Infrastructure (AMI) in WDNs with the use of smart meters and networking technologies that allow the transmission of data to a centralized system and track water usage in real time allows water operators to monitor demand, and identify leaks and other issues more quickly (Tuser, 2022).

DA in WDNs

Todini (1999) first introduced the application of Kalman Filter (KF) for calibrating pipe roughness coefficients in Water Distribution Networks (WDNs) with a linear structure. To handle nonlinear systems, Shang et al. (2008) utilized Extended Kalman Filter (EKF) to estimate nodal demands in a small hypothetical network by approximating nonlinearities with tangent linear operators. While these studies demonstrated good results with KF and EKF in cases of limited nonlinearity and uncertainty, their effectiveness may be hindered in highly looped networks (Van den Bossche, 2012) or when dealing with significant measurement errors (Shang et al., 2008).

The effectiveness of Ensemble Kalman Filter (EnKF) was demonstrated by Okeya et al. (2014a)in updating water demands and demand model parameters for a Water Demand Forecasting Model, assuming known pipe roughness values and no leakage. Okeya et al. (2014b) extended this approach to burst detection by applying Kalman filtering to flow observations and forecasts from a hydraulic model, and later, Okeya et al. (2015) showed that this methodology effectively detected bursts in real-time and estimated leak flow. Ruzza (2017) conducted a similar leak detection study in WDNs using KF, EnKF, Ensemble Smoothing, and Normal-Score EnKF for identifying nodal leakages. Zhou et al. (2018, 2022). demonstrated that ensemble-based methods ensure stable calibration results and long-term accuracy of models.

EnKF avoids the need for model linearization by simulating model states using an ensemble of parameters obtained through Monte Carlo perturbations. Particle Filter (PF), which extends the use of the ensemble to non-Gaussian models and increases the ensemble size, was successfully employed by Do et al. (2017) to estimate nodal demand patterns in WDN models, accounting for measurement errors. Bragalli et al. (2016) conducted a recent study testing the use of EnKF in WDNs, employing an innovative 3-step EnKF approach for a small WDN, which displayed promising results in terms of multi-step data assimilation capabilities.

EnKF is an ideal and optimal method for applying DA for WDN as EnKF is stable with large nonlinear systems and a low probability of divergence from the true value. The computational demand of EnKF is also lower than PF (Gillijns et al., 2006; Simon, 2006; Van den Bossche, 2012). A comparison of various DA methods with EnKF is provided in Table 1.

Table 1

Characteristics of DA Methods

DA Method	Applicable Systems	Applicable System Size	Divergence	Computational Demand
KF	Linear	Small	Low	Low
EKF	Non-Linear (First order)	Small	High	Medium
PF	Non-Linear	Large	Low	High
EnKF	Non-Linear	Large	Low	Medium

Source: Gillijns et al., 2006; Simon, 2006; Van den Bossche, 2012.

In summary, there is a need for accurate and efficient modelling of WDNs. To improve the design and rehabilitation works of WDNs such as optimizing system performance via calibration, improving water supply reliability, and improving operational decision-making by enabling the identification of anomalies, preventing leaks and reducing operational costs. Current models do not fully capture the complexity and variability of real-world systems, leading to suboptimal decision-making. In this research, we explored the potential of the proposed DA technique to reduce model error, explore the factors that influence the effectiveness of the applied DA and assess its practical applicability.

Case Studies

One theoretical network which is representative of a real-world WDN is taken for this study, Modena network which is the same WDN used by many similar studies. Bragalli et al. (2016) utilized the WDN to understand the effectiveness of a 3-step EnKF DA method, Han et al. (2020) utilized it for seismic resilience enhancement of urban water distribution system and Bhave and Gupta (2006) used in for water quality based reliability analysis. The network consists of 317 pipes, 268 nodes and 4 reservoirs with a fixed head between 72.0 m and 74.5 m. The layout of the network is given in Figure 1. The network has a total length of 71. 8 km of pipes with diameters between 100 mm and 400 mm. In the upcoming section we will delve into the methodologies used to test the effectiveness of the DA Algorithm with the Modena network.

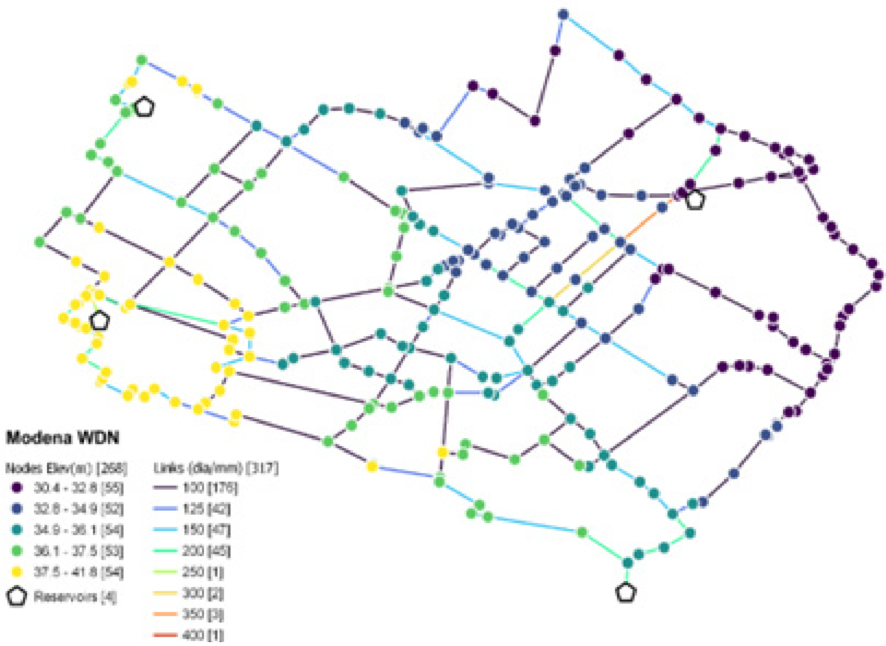

Figure 1: Modena WDN

Methodology

The EnKF DA algorithm was used for the DA. More details on the DA can be found in Fayaz et al (2023, under preparation).

Evaluation Metric

As a measure of the effectiveness of the system, the Total Variance (TV) is calculated similarly to Bragalli et al., (2016). To utilize the TV for extended period simulation the daily average TV value was obtained. Also, to utilize the TV for optimization the value was normalized the sum of the normalized average TV of all system states were summed and the normalized average was obtained.

Total Variance

$$TV\{\overline{\otimes}\} = \frac{1}{S}\sum_{i=1}^{S}\left(\overline{\otimes}_i - \overline{\otimes}_i^{true}\right)^2 + \frac{1}{S}\sum_{i=1}^{S}\left[\frac{1}{m(m-1)}\sum_{j=1}^{m}\ (\otimes_i^j - \overline{\otimes}_i)\right]^2$$

eq. 3-1

TV = Total Variance

$\otimes$ = State Variable (H, Q or q)

$\overline{\otimes}$ = Ensemble mean

S = number of state variables (number of nodes or pipes)

m = number of ensembles

i = iterator for the state variable

j = iterator for ensembles

Applying the DA for Optimization of Sensor Locations

The location of sensors plays an important part in understanding the hydraulic state of the WDN. The location of sensors determines the quality and sometimes the quantity of data that can be collected which impacts the efficiency and accuracy of the DA Algorithm. In some cases, there may a limited number of sensors available to be installed due to budgetary or time constraints. As such sensors need to be strategically placed in critical locations that can provide the most amount of information and allow the most significant reduction of uncertainty through the utilization of the DA algorithm.

In this study, an optimization problem was defined to find the optimal configuration of sensors. A Greedy Algorithm (GA) was utilized to find the optimal configuration of sensors with the objective function of minimizing the sum of the normalized average TV. GA works by making locally optimal solutions at each step incrementally improving the solution at each step, leveraging on the current best result, in the hope of obtaining the best overall solution.

The GA works with the following steps for each sensor type:

1. A sensor location is selected from the WDN. A node for head or demand and a link for flow sensors
2. Run the DA Algorithm and calculate the sum of the normalized average TV based on the outputs of the DA Algorithm
3. Iterate over the remaining locations of sensors which have not been selected and repeat step 2 for each location.
4. Select the location that has the lowest sum of the normalized average TV and add it to the list of configured sensors.
5. Repeat steps 2 – 4 until the desired number of sensors is reached.
6. After all sensor types have been configured, the final sensor configuration of head, demand and flow sensors can be obtained.

It is important to note that the GA does not guarantee a globally optimal solution, but literature shows that GA is effective in obtaining near-optimal solutions within a relatively short period as the search space consecutively gets smaller as each sensor gets optimized.

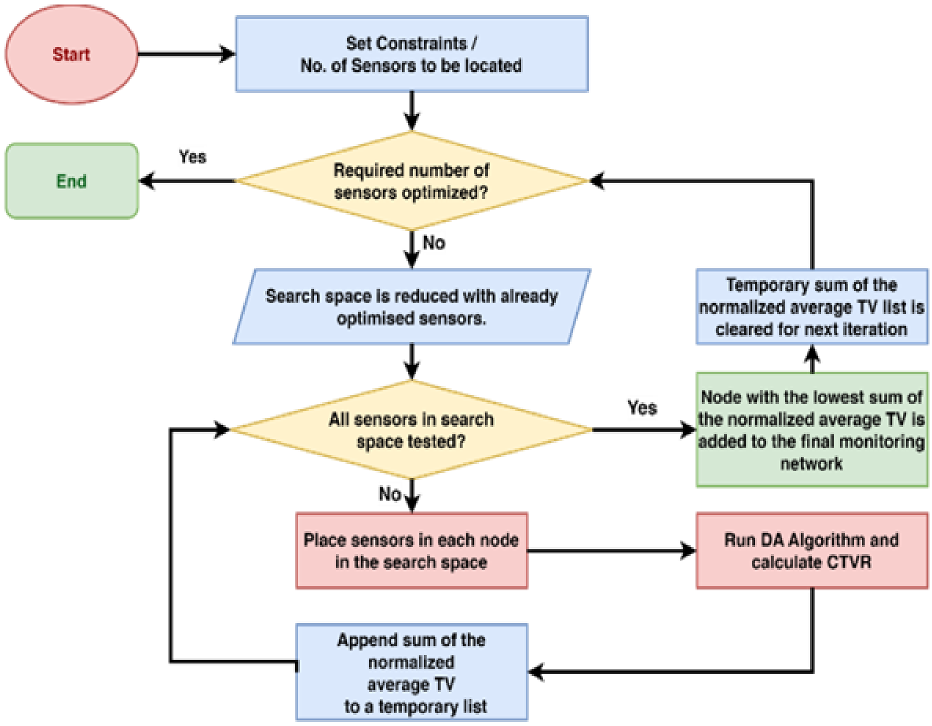

Figure 2: Framework of the GA used for optimizing sensor locations for DA

Results and Discussion

The DA algorithm assimilates head measurements followed by flow measurements and lastly demand measurements. The results of the optimization of the sensor for Modena are shown in Figure 3 and Figure 4. As seen from the results, the increase in the number of sensors successfully decreases the sum of the normalized average TV of the DA showing a consecutive improvement in the results as the number of sensors is increased. Due to the significantly high computational demand, only up to 40 sensors were optimized for this study. However, an asymptotic limit can be seen within the number of sensors optimized, where an increase in sensors beyond 28 – 30 sensors does not show a significant improvement in the effectiveness of the DA in the case of Modena.

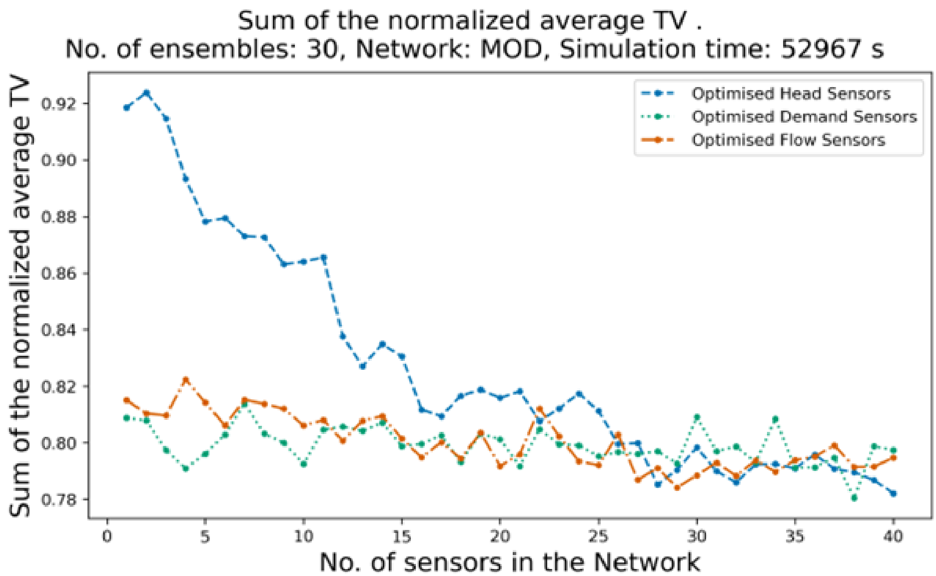

Figure 3: Sum of normalized average TV against no. of optimized sensors for Modena

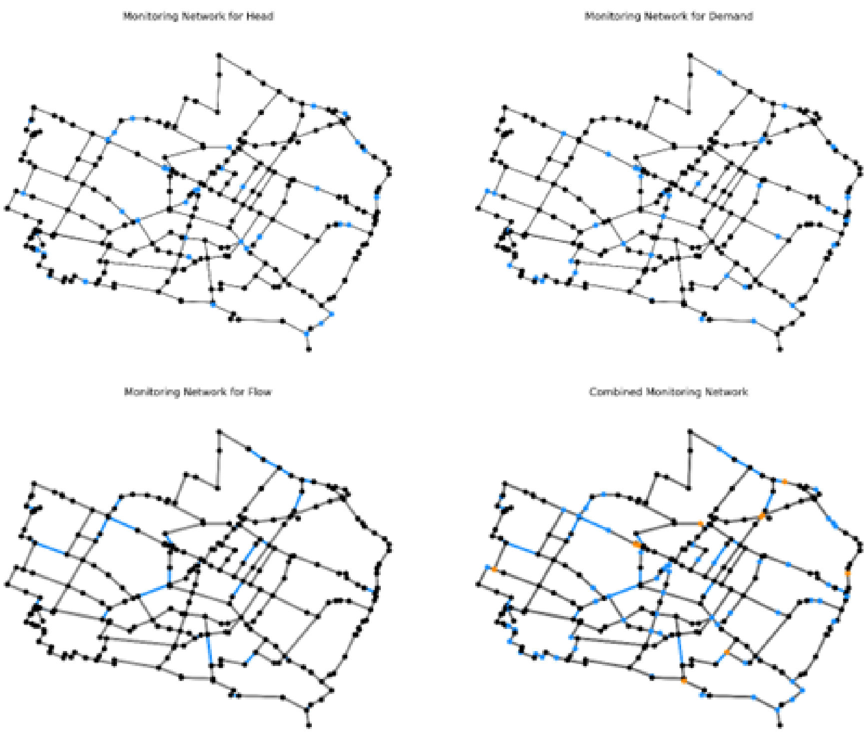

Figure 4: Optimized Monitoring Networks for Modena WDN (Blue: One Sensor, Orange: 2 Sensors)

Conclusions and Recommendations

The main objective of this research study was to examine the efficacy of using DA for reducing the WDN model error and exploring the potential of using DA to locator sensors optimally in the WDN.

The results showed that an increase in the number of sensors of each type does improve the effectiveness of the DA. However, despite the use of a GA, the computational demand of the DA for an iterative search of single sensors proved to be quite demanding. It was also seen that beyond 28-30 sensors, does not significantly improve the results of the DA, suggesting additional sensors not providing a significant improvement in the state estimation of the WDN. Hence, it is understandable that water operators can opt to optimize the use of sensors in their WDN by using the proposed DA and GA algorithms to ensure an optimized number of sensors are used in their respective WDN's, reducing the investment costs of different types of sensors needed for the WDN.

The use of the DA in WDN shows promising results and implementation of a multi-step DA has advantages that allow the use of all available information. The GA optimization also shows that depending on the WDN characteristics a limited number of optimized sensors can provide a significant improvement in the error in the WDN models, without extensive monitoring of the WDN with unoptimized sensors. Further development of the 3-step DA to include detection of leaks and anomalies, calibration of WDN models can also be explored.

Conflict of Interest

We hereby declare that we have no conflict of interest related to this research.

Funding

This research was supported by the internship program of Vitens N.V.

Acknowledgements

The authors wish to acknowledge and appreciate the support and contribution provided by Vitens N.V.

References

Antonowicz, A., Brodziak, R., Bylka, J., Mazurkiewicz, J., Wojtecki, S., & Zakrzewski, P. (2018). Use of EPANET solver to manage water distribution in Smart City. E3S Web of Conferences, 30, 01016. https://doi.org/10.1051/e3sconf/20183001016

Bhave, P. R., & Gupta, R. (Rajesh). (2006). Analysis of water distribution networks. Alpha Science International; WorldCat.org.

Bragalli, C., Fortini, M., & Todini, E. (2016). Enhancing Knowledge in Water Distribution Networks via Data Assimilation. Water Resources Management, 30(11), 3689–3706. https://doi.org/10.1007/s11269-016-1372-0

Do, N. C., Simpson, A. R., Deuerlein, J. W., & Piller, O. (2017). Particle Filter–Based Model for Online Estimation of Demand Multipliers in Water

Distribution Systems under Uncertainty. Journal of Water Resources Planning and Management, 143(11), 04017065. https://doi.org/10.1061/(asce)wr.1943-5452.0000841

ECMWF. (2013, November 29). Data assimilation [Text]. ECMWF. https://www.ecmwf.int/en/research/data-assimilation

Gillijns, S., Barrero Mendoza, O., Chandrasekar, J., De Moor, B., Bernstein, D. S., & Ridley, A. (2006). What is the ensemble Kalman filter and how well does it work? 1–12, 6 pp. https://doi.org/10.1109/ACC.2006.1657419

Han, Z., Ma, D., Hou, B., & Wang, W. (2020). Seismic Resilience Enhancement of Urban Water Distribution System Using Restoration Priority of Pipeline Damages. Sustainability, 12(3), 914. https://doi.org/10.3390/su12030914

Mathworks. (2022, September 24). Understanding Kalman Filters, Part 3: Optimal State Estimator Video. Mathworks. https://nl.mathworks.com/videos/understanding-kalman-filters-part-3-optimal-state-estimator--1490710645421.html

Okeya, I., Hutton, C., & Kapelan, Z. (2015). Locating Pipe Bursts in a District Metered Area Via Online Hydraulic Modelling. Procedia Engineering, 119, 101–110. https://doi.org/10.1016/j.proeng.2015.08.859

Okeya, I., Kapelan, Z., Hutton, C., & Naga, D. (2014a). Online Burst Detection in a Water Distribution System Using the Kalman Filter and Hydraulic Modelling. Procedia Engineering, 89, 418–427. https://doi.org/10.1016/j.proeng.2014.11.207

Okeya, I., Kapelan, Z., Hutton, C., & Naga, D. (2014b). Online Modelling of Water Distribution System Using Data Assimilation. Procedia Engineering, 70, 1261–1270. https://doi.org/10.1016/j.proeng.2014.02.139

Rossman, L. A. (1994). EPANET users manual: Project summary. Cincinnati, OH : U.S. Environmental Protection Agency, Risk Reduction Engineering Laboratory, [1994]. https://search.library.wisc.edu/catalog/999778236102121

Ruzza, V. (2017). Data Assimilation Techniques for Leakage Detection in Water Distribution Systems. Universit`a degli Studi di Padova.

Shang, F., Uber, J. G., van Bloemen Waanders, B. G., Boccelli, D., & Janke, R. (2008). Real Time Water Demand Estimation in Water Distribution System. Water Distribution Systems Analysis Symposium 2006, 1–14. https://doi.org/10.1061/40941(247)95

Simon, D. (2006). Optimal state estimation: Kalman, H [infinity] and nonlinear approaches. Wiley-Interscience. https://isharifi.ir/teaching/2019/IoT/[Dan_Simon]_Optimal_State_Estimation_Kalman,_H_In(BookFi).pdf

Todini, E. (1999). Using a Kalman filter approach for looped water distribution network calibration. In Water Industry systellls: Modelling and Optimization (Vol. 1). Research Studies Press.

Tuser, C. (2022, March 1). What is Advanced Metering Infrastructure (AMI)? | Water & Wastes Digest. Water and Wastes Digest. https://www.wwdmag.com/editorial-topical/what-is-articles/article/10940067/what-is-advanced-metering-infrastructure-ami

United Nations. (2022a, September 23). Goal 6 | Department of Economic and Social Affairs. Ensure Availability and Sustainable Management of Water and Sanitation for All. https://sdgs.un.org/goals/goal6

United Nations. (2022b, December). Water. United Nations; United Nations. https://www.un.org/en/global-issues/water

University of Reading. (2022, September 24). What is data assimilation? [University]. DARC. https://research.reading.ac.uk/met-darc/aboutus/what-is-data-assimilation/

Van den Bossche, W. (2012). Wannes Van Den Bossche—Data Assimilation Toolbox For Matlab [KU Leuven]. https://www.scribd.com/document/409269852/Wannes-Van-den-Bossche-Data-assimilation-toolbox-for-Matlab-pdf

Zhou, X., Guo, S., Xin, K., Xu, W., Tao, T., & Yan, H. (2022). Maintaining the long-term accuracy of water distribution models with data assimilation methods: A comparative study. Water Research, 226, 119268. https://doi.org/10.1016/j.watres.2022.119268

Zhou, X., Xu, W., Xin, K., Yan, H., & Tao, T. (2018). Self-Adaptive Calibration of Real Time Demand and Roughness of Water Distribution Systems. Water Resources Research, 54(8), 5536–5550. https://doi.org/10.1029/2017WR022147