MSc Thesis

# Leak Localization in Water Distribution Networks

Improvement of a model-based method and exploration of a data-driven method for leak localization using pressure sensors

Zixi Meng

Dr. Riccardo Taormina (Chair), Prof. Dr. Zoran Kapelan, Dr. Nazli Yonca Aydin (TU Delft)

Dr. Martijn Bakker (Royal HaskoningDHV)

Water Management

Department of Civil Engineering and Geosciences, TU Delft

# Abstract

Leakage is the main source of water loss in water distribution networks (WDNs). Therefore, leak detection and localization technology is a major concern for water utilities to save water and meet the ever-growing water demand. This study presents two methodologies for leak localization in District Metered Areas (DMAs): (1) a model-based method using the hydraulic model, flow and pressure measurements, as well as leak flow; (2) a data-driven method that relies on graph-based interpolation. The performance of the model-based method is proved to be negatively affected by model errors and limited sensors. To solve these two problems, on the one hand, flows and residuals between observed and model-simulated data in the non-leak situation are used to develop a residual model to calculate offset values for model output correction. On the other hand, graph-based interpolation is introduced to create 'virtual' sensor measurements in the presence of a limited number of sensors. The data-driven method proposed in this work uses graph-based interpolation to estimate the head signals at the nodes without sensors and subsequently create pressure maps. Leak localization is achieved by comparing pressure maps in the non-leak and leak situations. In this process, this methodology does not require a well-calibrated model and leak flow information. Both two methodologies are tested on fire hydrant leak tests in DMAs in the United Kingdom. Results obtained by using the model-based method illustrate the positive impact of model output correction on localization results and the performance of this method under different conditions such as different times of the leak and different sizes of leak flow. The data-driven method performs fairly well in DMAs with a higher spatial density of sensors. Furthermore, the results of both two methods are compared to demonstrate the suitability of the methods in different cases.



**Keywords:** Leak localization; Model-based methods; Data-driven methods; Graph-based interpolation; Water distribution networks

# Contents

# Chapter 1. Introduction

## 1.1 Background Introduction

Water is a precious resource on our planet and plays an important role in human daily life. Although water covers around 71% of the earth, only a limited amount of water is drinkable and easily accessible. Water scarcity becomes an increasingly serious problem as climate change intensifies and the population grows. According to the report from WHO in 2019, approximately 10% of people globally are unable to access safe drinking water and half of the world does not have safe sanitation services (WHO, 2019).

Drinking water is generally delivered to users through water distribution networks (WDNs). In this process, a significant challenge for water companies is the high level of water loss (difference between the volume of water put into the water distribution system and the volume charged to customers), which is also known as Non-Revenue Water (NRW) (Liemberger & Wyatt, 2019a). The main cause of water loss in WDNs is leakage, accounting for almost 70% of the total water loss (El-Zahab & Zayed, 2019). The volume of water leakage in WDNs varies considerably between different countries, ranging from 3% to 7% in the WDNs in good conditions (e.g., the Netherlands) to more than 50% in certain developing countries as shown in Figure 1 (Puust et al., 2010). In addition to the water quantity issue, leakage also leads to economic loss, environmental contamination, and potential health problems (Puust et al., 2010). An estimation of 126 billion cubic meters of NRW per year worldwide equals about 39 billion USD lost per year (Liemberger & Wyatt, 2019b). In 2015, engineers at the University of Sheffield proved that potentially harmful contaminants can enter through leaky water pipes and be transported through the WDN (University of Sheffield, 2015).

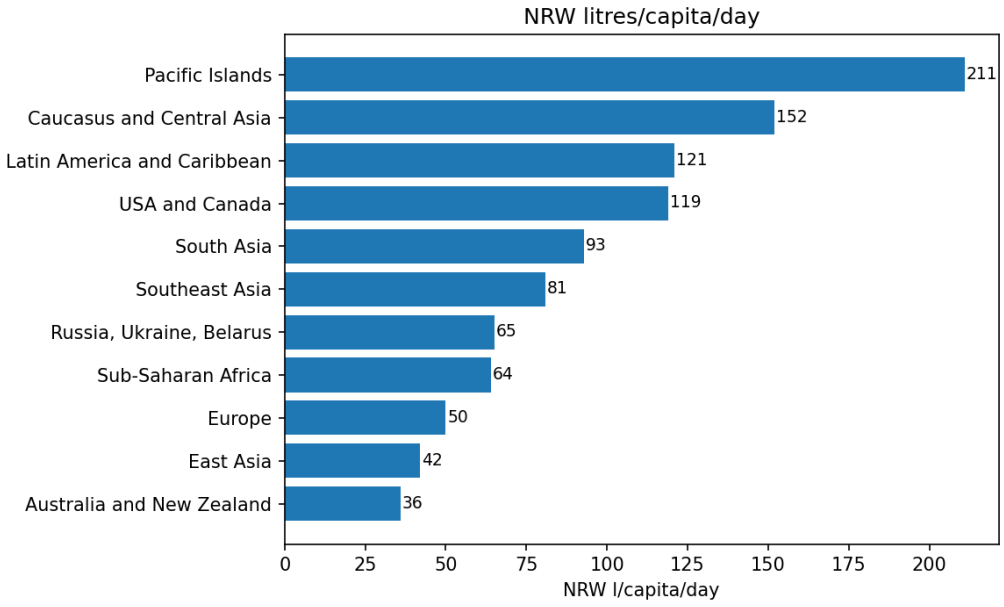


**Figure 1.** NRW in different regions (Liemberger & Wyatt, 2019b).

Figure 2 shows five phases involved in a burst cycle. Burst duration is determined by the overall period between the occurrence and reparation of the burst (Bakker, Vreeburg, et al., 2014a). Typically, large bursts are visible and easy to be detected and reported to water companies via the internet or service lines. Small leaks underground

often remain unnoticed for a long time, resulting in greater water loss (Wu & Liu, 2017). Therefore, when small leaks occur, it is crucial that water company operators are quickly aware of leaks in WDNs and then they provide timely information to maintenance technicians using leak detection and localization approaches, thereby helping them to find the exact leak and fix it as soon as possible. However, small leaks are quite challenging because they cause only weak hydraulic disturbances in WDNs, which can easily be masked by other unexpected behaviors in WDNs. This study will focus on leak localization techniques based on analyzing the hydraulic impact of the leak to shorten the 'Location period' (L) and save water.

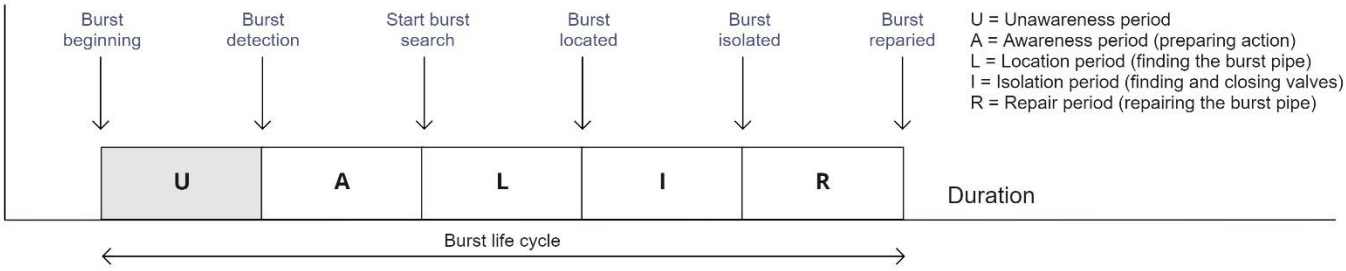


**Figure 2.** The life cycle of a burst event (Bakker, Vreeburg, et al., 2014a).

## 1.2 Literature review

Various methods were studied for leak detection and localization, and they can be divided into hardware-based and software-based methods (Li et al., 2015; Sophocleous et al., 2019; Zaman et al., 2020) as reported in Figure 3 below.

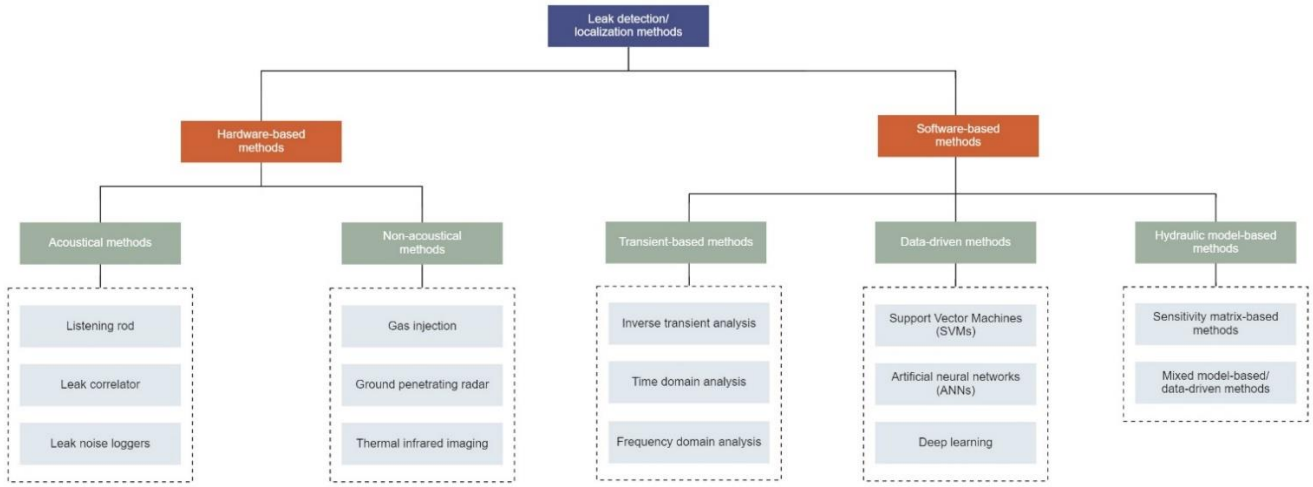


**Figure 3.** Classification of leak detection/localization methods (Alves et al., 2022; R. Li et al., 2015).

Hardware-based approaches are commonly used for leak localization by water utilities, including acoustic emission methods such as listening rods, leak correlators as well as noise loggers, and non-acoustic methods like thermal infrared imaging (R. Li et al., 2015). Acoustic methods are less effective in the case of large diameters and new pipe materials (for example polyvinyl chloride, PVC, which is used in over half of the drinking water networks in the Netherlands), where the leakage sound is low in frequency and difficult to identify (Gao et al., 2005;

Sophocleous et al., 2019). Hardware-based methods are costly since they are labor-intensive and sometimes require expensive equipment. Software-based methods use in-situ sensors to monitor continuous data, such as pressure and flow instead of leak noise information between simulated and measured differences, to estimate the leak location using algorithms or models, making them well-adapted to diverse pipes. These approaches are more efficient and effective in reducing the search area of leaks. The combination of software-based and hardware-based methods makes it possible to pinpoint the leak eventually in a more timely and accurate manner.

Software-based methods can be further divided into transient-based, model-based, and data-driven methods (Alves et al., 2022). Colombo et al. (2009) reviewed the transient-based leak detection methods, including inverse transient analysis, time domain analysis, and frequency domain analysis, which rely on transient flow analysis to extract high-frequency data from the generated pressure waves (R. Li et al., 2015). These approaches are mainly used on single pipelines above the ground due to the high effect of the system configuration and uncertainty (Puust et al., 2010). Furthermore, a vast number of sensors and complex computations are required in this process, resulting in extremely high costs.

Model-based approaches use the hydraulic model to simulate the leak in the model environment and identify the leak by comparing the leak-driven hydraulic disturbance in the model system and the real networks. The most common model-based method is to consider leak localization as a fault isolation problem by doing pressure sensitivity analysis (Casillas Ponce et al., 2014; Perez et al., 2014; Pérez et al., 2011; Quevedo et al., 2011). Farley et al. (2013) used Genetic Algorithms (GA)-based optimization of the Jacobian sensitivity matrix to reduce the search area by subdividing a DMA. Correlation is calculated between the leak-driven pressure drops in the model environment and the real world. Perez et al. (2014) used the model-based method with non-binary pressure sensitivity analysis for leak localization and achieved satisfactory results on a 30-h fire hydrant leak test in DMA Nova Icaria in Spain. Although model-based methods have been validated in both simulated and real-life applications, their performance is strongly affected by model errors, measurement noise, and unknown nodal demands (Blesa & Pérez, 2018). To account for these uncertainties involved, mixed model-based/data-driven methods that combine model simulation and machine learning classification are proposed, such as K-nearest neighbors (KNNs) classifier (Soldevila et al., 2016) and Bayesian classifier (Soldevila et al., 2017), and Support Vector Machines (SVM) (Zhang et al., 2016).

Data-driven approaches use data measurements coupled with algorithms to find the leak, which are less dependent on hydraulic model. Quiñones-Grueiro et al. (2018) developed an approach using SVMs to interpret pressure data for leak flow estimation and leak localization. Other research on machine learning algorithms, such as artificial neural networks and clustering has been conducted as followed (Romano et al., 2014; Wu et al., 2016). Next, deep learning algorithms like Convolutional Neural Networks (CNNs) have been applied for leak localization due to their strong capabilities in feature identification and pattern recognition. Javadiha et al. (2019) used off-line pressure measurements and the model simulated pressure in the non-leak situation to create a pressure residual map with the help of Kriging interpolation. These residual maps were subsequently converted to 2-D pictures as image data corresponding to each leak location. Afterward, the authors proposed a CNNs solution for providing image classification and tested it on a benchmark network in Hanoi, Vietnam. J. Li et al. (2022) used a new densely connected convolutional network (ResNet) to improve the accuracy of leak localization. Data-driven methods usually require high quality and a large quantity of monitored data. Nonetheless, it is challenging to obtain valid data over a long time, especially for the data in different leak scenarios. Even if the long-term data is available, these methods are more operative for identifying the leak event (leak flow) that has previously occurred in WDNs. Apart from AI-based learning algorithms, Soldevila et al. (2019) and Soldevila et al. (2021)

compared the pressure at each node in non-leak and leak situations by using Kriging interpolation to locate the leak, which only requires historical pressure data in normal situations.

## 1.3 Problem statement

Although an increasing number of studies on promising software-based methods have been done, they are far from ideal. These techniques are not commonly applied by water utilities in real cases due to underdetermined performance of existing model-based and data-driven method (Romano et al., 2013; Zaman et al., 2020). This work proposes two different methods to locate the leak. Both of them are applied to real cases, using them to improve the localization performance.

The first method builds on the model-based method developed by Quevedo et al. (2011) using a hydraulic model and pressure and flow measurements. Despite the test and validation in real cases, this method is not widely used by water utilities since the performance is limited to the accuracy of the WDN model. Some model information such as pipe roughness coefficients and nodal demands is difficult to precisely measure, resulting in deviations between model outputs and measurements, which can lead to incorrect localization. Traditional solution to solve this problem is model calibration (Savic et al., 2009). However, it is costly for either computation or manpower since a lot of adjustment work on model parameters is required to meet the calibration criteria(Preis et al., 2011; Sophocleous et al., 2019). Moreover, the limited number of sensors is another issue making it more difficult to capture the pressure changes driven by the leak in WDNs. This work presents a more efficient solution to compensate for the model errors in more DMAs at once to improve the leak localization, which has better prospects for commercial application. In addition, data interpolation is introduced to provide more available data (Soldevila et al., 2019).

The other method is a data-driven method using pressure data as well as interpolation, which is an alternative method to achieve accurate leak localization without well-calibrated models. The basic idea comes from Soldevila et al. (2019) and Soldevila et al. (2021), in which the authors compare the pressure at each node (pressure map) in the non-leak and leak situations, where the pressure map is obtained by Kriging interpolation. However, Kriging interpolation does not consider the topological structure and hydraulic features of WDN. In this study, graph-based interpolation is introduced due to its high better performance using synthetic data (X. Zhou et al., 2022a). It is explored for the first time using real data and employed for leak localization in real cases.

Furthermore, methodologies for leak localization are rarely tested and validated in real cases, especially data-driven methods. Given the influence of different factors on leak localization, the performance of both these two methods is investigated in different types of real DMAs and evaluated by two new performance metrics to provide an understanding of the applicability of these two methods in real life.

## 1.4 Research questions

The research objectives are to improve the model-based method and data-driven method for leak localization to achieve better localization performance and to test these methods in different situations in real DMAs. The following questions with specific sub-questions are formulated as follows:

Q1. How can we improve the model-based leak localization relative to the method developed by Quevedo et al. (2011)?

SubQ1-1. How to improve the model-based leak localization by minimizing the residuals between observed and model-simulated data?

SubQ1-2. Does graph-based interpolation improve localization precision?

SubQ1-3. How does the localization performance vary under different conditions such as different structure of networks and different leak flow?

Q2. Can we advise a data-driven approach using pressure and interpolation for leak localization in real WDNs?

SubQ2-1. How does the graph-based interpolation perform in real WDNs?

SubQ2-2. How does this data-driven approach perform in real WDNs?

SubQ2-3. Compared with the model-based method, in which cases the data-driven approach is recommended for practical use?

## 1.5 Thesis structure

To answer the research questions in section 1.4, Chapter 2 will first show the framework of the research, explaining the relationship between the research questions and each chapter. Next, the model-based and data-driven methods (including interpolation) will be explained respectively. Chapter 3 will show the basic information of the case study used for this research and describes the plan for artificial leak tests. Chapter 4 will present the localization accuracy of both model-based and data-driven approaches and discuss the results. Chapter 5 is the conclusion of this thesis and provides recommendations for future studies.

# Chapter 2: Methodology

Figure 4 illustrates the overall framework and flow of this research. The leak detection is out of the scope of this work, and it is assumed that the leak has been detected and the leak flow has been estimated efficiently in the DMA. The methods are used to find the leak which is simulated by opening the fire hydrant. The leak size is fixed as planned during the leak test and the leak location is determined by technicians from the water company. The methodology involves two parts: (1) a model-based method (in section 2.1) and (2) a data-driven method (in section 2.3) for leak localization. Hydraulic simulations were carried out using the software EPANET 2.2 and the WNTR package in Python (Klise et al., 2017). In the first part, a residual model is proposed to simulate the deviation between model outputs and measurements in the non-leak situation. Next, graph-based interpolation is integrated into the model-based method. In the second part, the accuracy of graph-based interpolation is first investigated using synthetic data in real WDNs to determine the variables used for leak localization. Then the data-driven method is implemented using real data for finding leaks in WDNs.

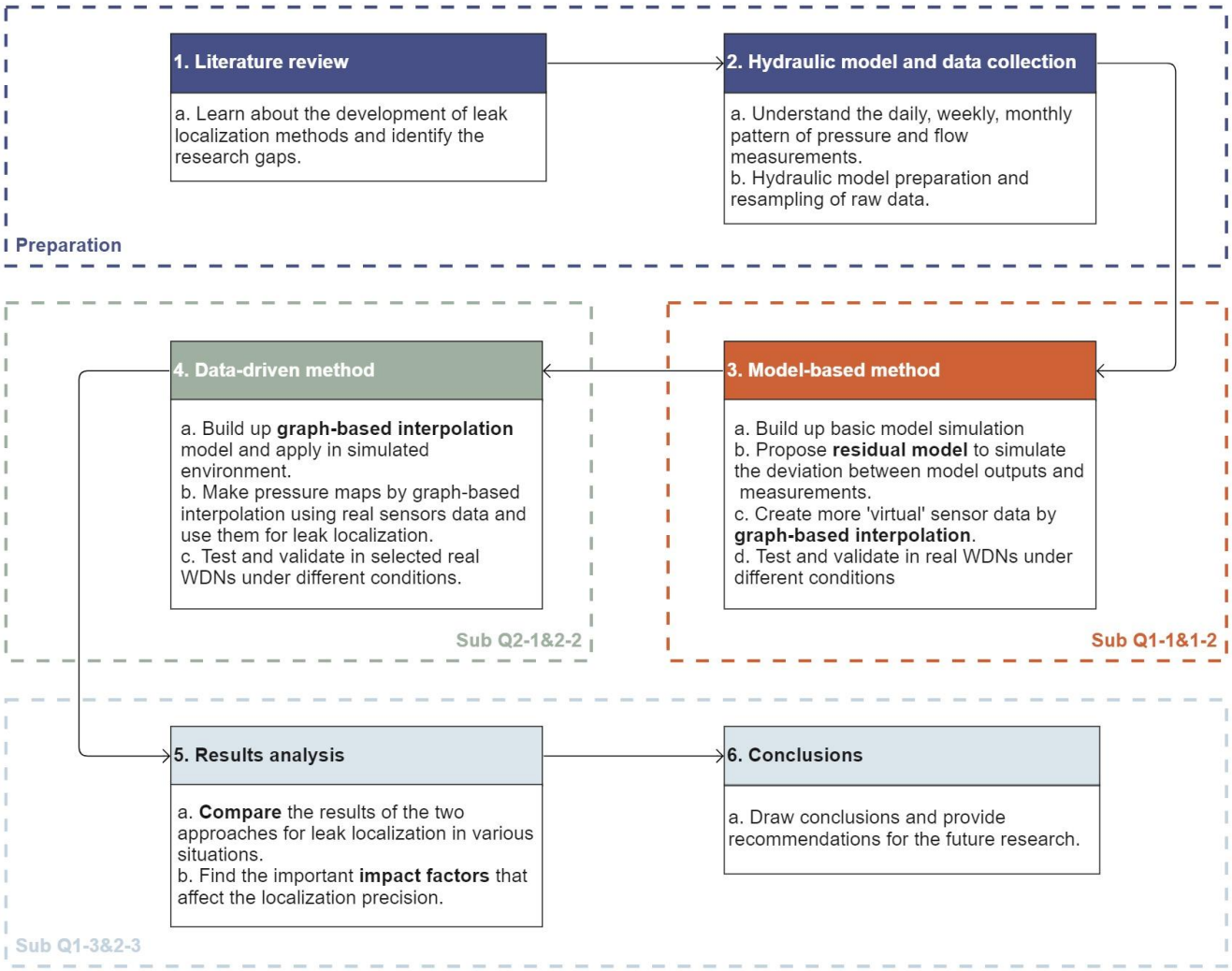


**Figure 4.** Overview of the research.

## 2.1 Model-based method using pressure sensors

This work aims at locating the leak in the simple DMA with a single inlet, which means no other elements are present in the DMA, such as outlets, pressure regulating valves (PRVs), and/or boosters. Flow measurements at the inlet, and pressure measurements at the inlet, critical point as well as additional five to eight nodes are available from Feb 2022. In this study, leak flow r is constant controlled by technicians during the time period of the leak test. It is assumed that only one leak occurs in the DMA (to simplify the problem) at a time (the same as the assumption in Casillas Ponce et al. (2014), Perez et al. (2014), Pérez et al. (2011),and Soldevila et al. (2019)).

### 2.1.1 General approach

In this work, the model-based method used for leak localization can be described as Figure 5, including three stages: (1) data acquisition, (2) preprocessing, and (3) leak localization. In the first stage, data of flow and pressure measurements is collected every 1 min. Next, a residual model is proposed for each sensor using the flow and residual between model outputs and measurements in non-leak situation. Faulty sensors are detected and discarded before localization analysis. After the leak has been detected in WDNs, flow during the leak is used to calculate the offset values to correct the model output using the residual model. Then, the model-based method relying on pressure sensitivity analysis is implemented to locate the leak. Finally, localization results are assessed by two novel performance metrics proposed in section 2.4.

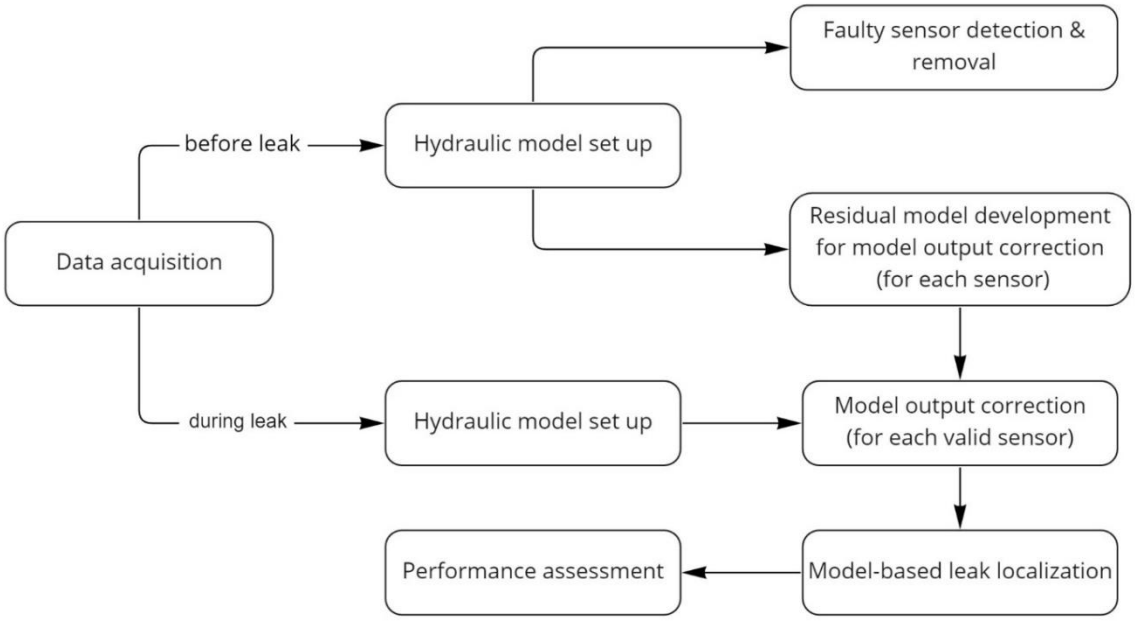


**Figure 5.** Schematic description of model-based method.

### 2.1.2 Hydraulic model set up

The hydraulic model of the entire service area was provided by the water company in the form of a geographic information system (GIS) file. The model of each DMA was split by shapefile from GIS and imported into EPANET 2.2. Network structure (including pipes and nodes) and parameters (such as pipe length, diameters, and

roughness) were retrieved from the GIS system. These models are supposed to be able to represent the behavior of WDNs. At the inlet of each DMA, the entry point is modeled as a water reservoir in EPANET. Leak is modeled as extra demand added to every node in the DMA. The setting of the hydraulic time step, pattern time step, and reporting time step are the same as the scenario analysis time step (which is further explained in section 2.1.5) equal to 15 min in the EPANET model.

The Darcy Weisbach formula is used for pipe head loss calculation in this study since it is the most theoretically correct one among all the head loss formulas (Rossman, 2000).

$$h_{fr} = f_D \frac{8}{\pi^2 g D^5} L Q^2 = k_{fr} Q^2 \quad (1)$$

where $h_{fr}$ is the pipe friction headloss; $f_D$ is the Darcy friction factor; $D$ is the pipe diameter; $L$ is the pipe length; $k_{fr}$ is the frictional loss coefficient; and $Q$ is the volumetric flow rate.

Boundary conditions flow and pressure are measured at the entry point of each DMA. Head (the summation of the pressure and elevation) is input and set as the head pattern for the modeled reservoir as shown in Figure 6. Flow measurement is used to scale the demand pattern. The scaling factor for each time step $t$ equals the flow divided by the sum of based demand of each node. Base demand is determined based on the billing information collected by water utilities and reflects the average consumption in the year prior to the year that the model was built.

$$Pattern\ multiplier(t) = \frac{Q(t)}{Sum\ of\ billed\ consumption\ all\ user} \quad (2)$$

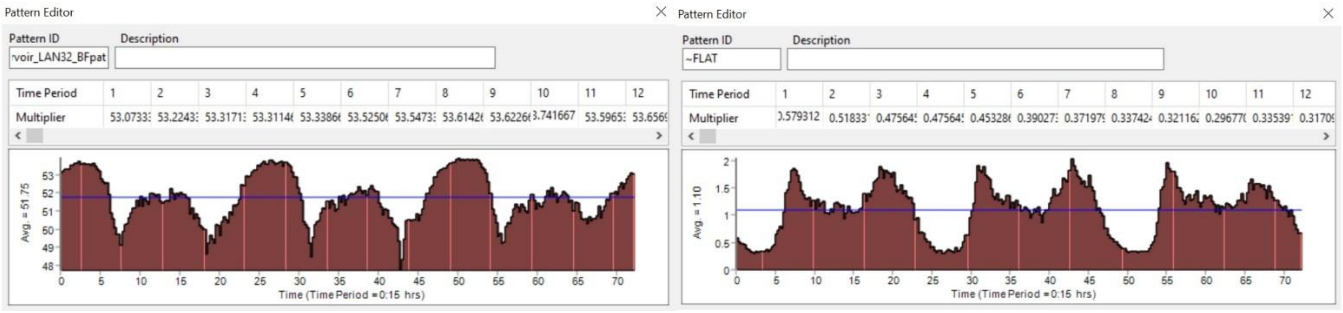


**Figure 6.** Flow and head patterns in EPANET. Left: Head pattern of the reservoir; Right: Demand pattern.

## 2.1.3 Residual model development

Nodal pressure is an essential indicator of hydraulic status of WDN (X. Zhou et al., 2022a). In an ideal situation, for each sensor, the pressure measurement should be the same as the model output in non-leak situation with perfect sensor and accurate mode. However, there are always residuals between measurements and model outputs (as shown in Figure 7) due to uncertainties particularly in nodal demand and pipe roughness coefficients as well as measurement noise (Soldevila et al., 2016). Here the measurement is the sum of measured pressure and elevation of a node. These residuals may lead to poor confidence in localization results (Perez et al., 2014).

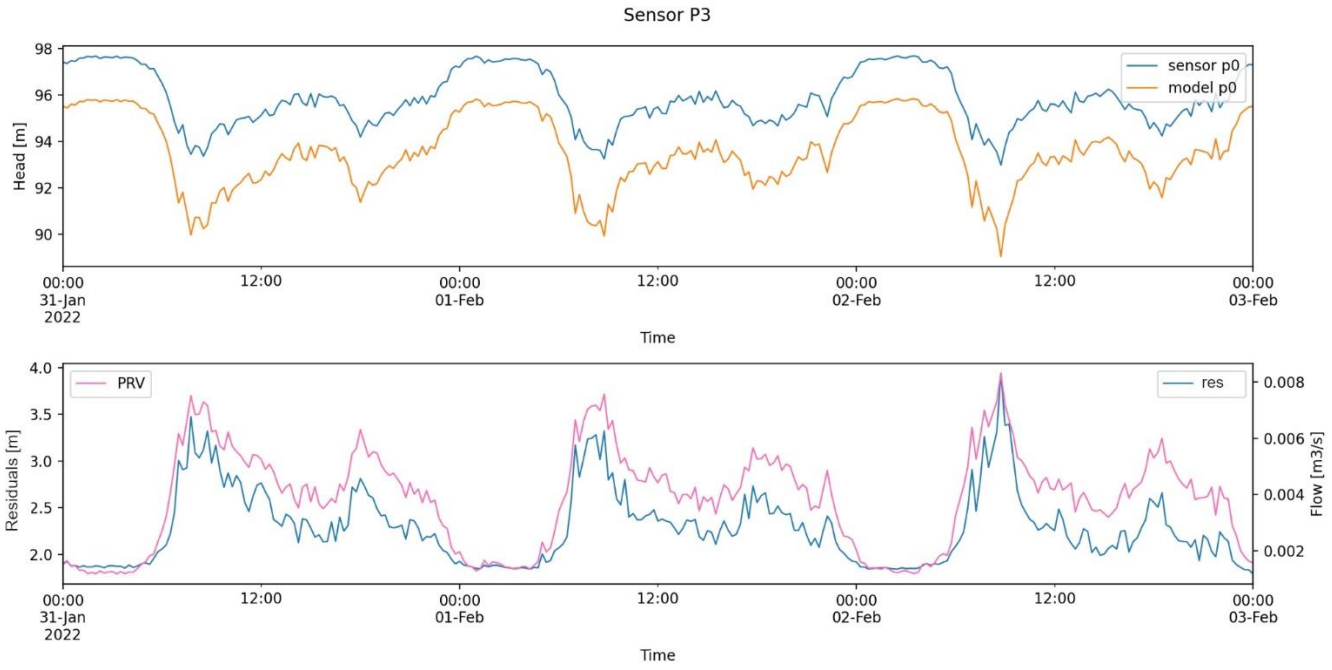


**Figure 7.** Residuals between observed and model-simulated pressure head. Upper: The observed (blue) and model simulated pressure head (orange) of sensor P3 in DMA5 in the non-leak situation. Lower: The inflow of DMA6 (pink) and residuals (blue).

In this study, for each node with sensor, we consider developing a simple model to simulate these residuals and to further use this model to calculate an offset to correct the model outputs. Residuals can be generally divided into static and dynamic errors. On the one hand, static errors are caused by elevation difference between sensor and the node in model. On the other hand, water flows from the entry point of the DMA to each sensor's location along pipes. Water head drops due to the pipe resistance in this process. However, the model cannot precisely simulate the head loss in real life because of underdetermined nodal demand and network parameters. In the model environment, head loss in pipes consist of both friction losses and local losses, which can be estimated as follow (Bentley Systems et al., 2003),

$$h_{loss} = h_{fr} + h_{minor} = (k_{fr} + k_{minor})Q^2 \quad (3)$$

where $h_{loss}$ is the total head loss; $h_{fr}$ is the pipe friction head loss; $h_{minor}$ is the minor loss; $k_{fr}$ is the coefficient of pipe friction head loss; $k_{minor}$ is the minor loss coefficient.

Based on eq (8) where the pressure head is quadratic related to flow in pipes, we assume that the residuals have the same relation with flow. Consequently, for each senor node i, a simple quadratic function is proposed to model the residuals.

$$R_i(t) = p_{meas,i}(t) - p_{model,i}(t) = C1_i Q^2(t) + C2_i \quad (4)$$

where $p_{meas}$ is the sum of measured pressure and elevation (at sensors i); $p_{model}$ is model simulated head; $C2$ is the dynamic correction factor; $C2$ is the static correction factor.

The parameters $C1$ and $C2$ are estimated by least squares approach using historical flow and residuals in non-leak situation. According to Eq. (4) and Figure 8, residuals highly depend on the flow varying with user behavior. They are higher and more variable during peak hours, and smaller and more stable at night. The same difference in

residuals also exists between working days and weekends due to different water consumption pattern (S. L. Zhou et al., 2002). Since the model is applied to calculate the offset to correct model outputs after leak, residuals and flows used to build the residual model should be similar to the WDN behavior during the leak. Considering these factors, the observed flows and residuals before the leak are divided into weekdays and weekends. In this study, if the leak is detected on a weekday, flows and residuals from 3 complete weekdays (72 hours) prior to the leak are selected to determine $C1$ and $C2$ for each sensor since it gives a good balance between residual model fit accuracy and the efficiency of computation (avoid too large dataset). If the leak is detected on a weekend, dataset from the weekends (48 hours) one week ago are selected. The total length of time steps depends on the scenario analysis time step $T_a$. The offset values used to correct model outputs during the leak are updated daily on weekdays and weekly on weekends.

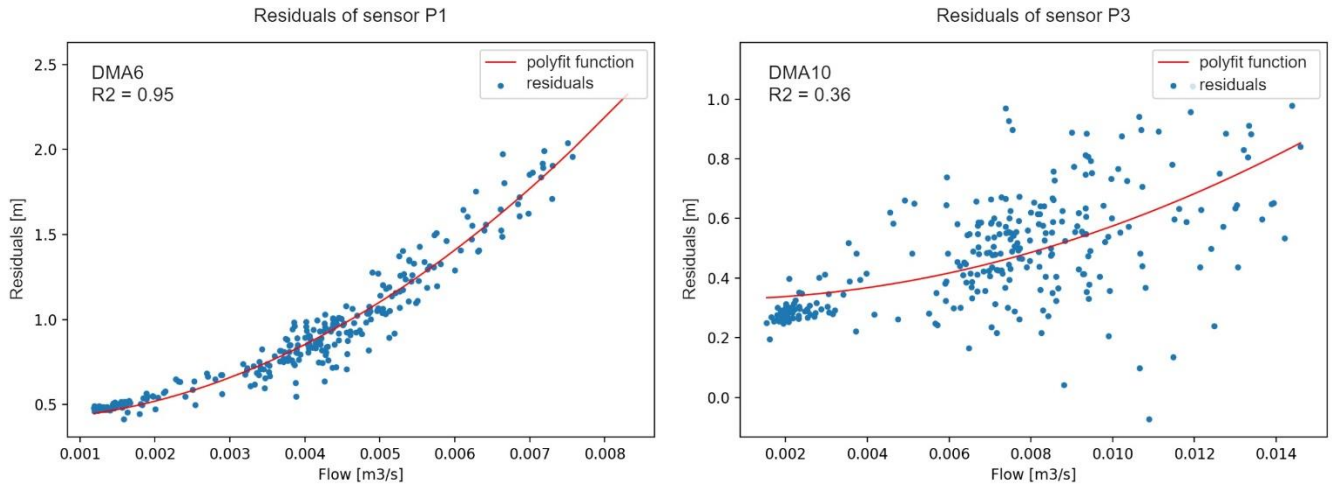


**Figure 8.** Fitting of the residual model.

## 2.1.4 Faulty sensor detection

Sensors are not always accurate since pressure signals are sensitive to changes in conditions. Faulty sensors behave abnormally, leading to different patterns of measurements and model outputs. Therefore, it is of interest to detect and exclude those faulty sensors to guarantee the localization precision (Perez et al., 2014). Figure 9 shows the flowchart of faulty sensor detection and removal. For the 72-h evaluation before the leak, each measurement is compared to model output value. Mean absolute percentage error (MAPE) is introduced to evaluate the differences between the reality and model at each sensor node i.

$$MAPE_i = \frac{1}{Np_{bnd}} \sum_{t=1}^{N} R_i(t) \times 100\% \quad (5)$$

where $N$ is the number of scenario analysis time step $T_a$; $p_{bnd}$ is the boundary head at the inlet of DMA.

If the MAPE > 5%, then the sensor is considered a faulty sensor and excluded before localization analysis. In other cases, model outputs are corrected by adding offset values calculated by residual model. The MAPE after model output correction (MAPE1) should be within $\pm$ 0.25% in the DMA since 0.25% noise is often added to perfect data to simulate real-life situations. Pressure response to a leak may be masked by large differences between measurements and model outputs.

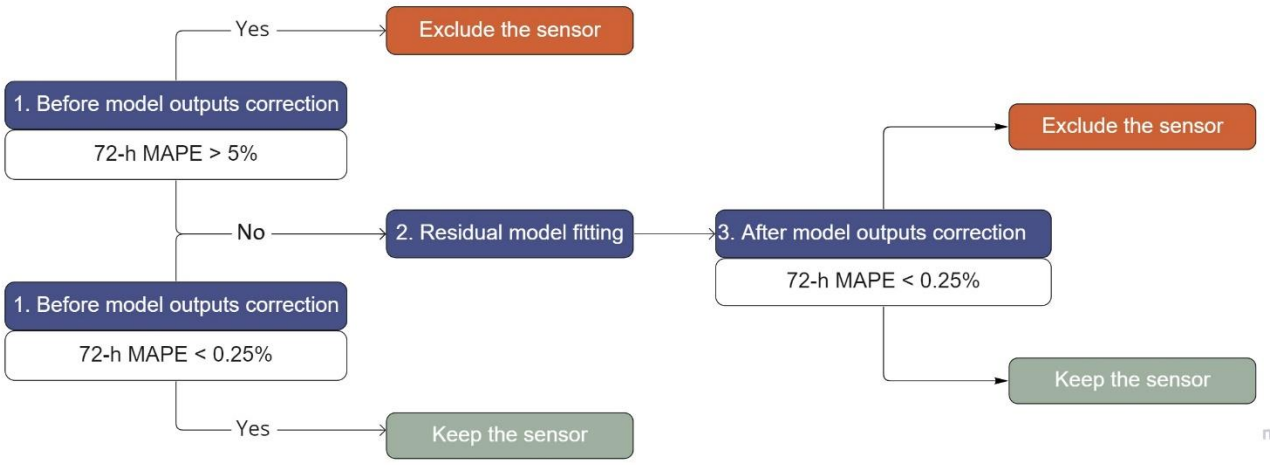


**Figure 9.** Flowchart of faulty sensor detection and removal.

### 2.1.5 Model-based leak localization

In Quevedo et al. (2011), leak isolation in WDNs is considered a fault isolation problem in the system. The main idea is of this model-based method to compute the correlation between the pressure drops driven by each potential leakage in the model network and the leak-driven pressure changes in real WDNs. The node with the highest correlation values is the most probable leaky node.

In this research, three different time steps are involved, including the sensor sample time $T_s$, scenario analysis time step $T_a$, and the diagnosis sliding window $T_w$ (Meseguer et al., 2014a; Moors et al., 2018). The $T_s$ of pressure sensors and flow meter is 1 min. The initial setting of $T_a$ is 15 min since it is the typical sampling rate of flow and pressure sensors (Bakker, Vreeburg, et al., 2014a; Romano et al., 2013). As a result, pressure and flow measurements should be first averaged over 15 minutes to reduce nuisance noise of measurement (Meseguer et al., 2014a). $T_w$ depends on the overall leak period, which in this study is the duration of artificial leak test.

The computation of correlation value for each node repeat at each $T_a$. First, the residual vector $RV$ is created to represent the pressure responses to the leak in real life. Hydraulic simulation is made to obtain the model outputs $p_0$ at the nodes equipped with pressure sensors without leak in the model. An offset value $p_{offset}$ calculated by residual model for each sensor node is added to $p_0$ to correct model outputs. Pressure measurements $p_i$ are acquired from each pressure sensor during leak event. Each item is calculated by subtracting corrected $p_0$ from $p_i$.

$$p_{offset,i}(t) = C1_i Q^2(t) + C2_i \tag{6}$$

where $p_{offset,i}$ is the offset value for model output correction at sensor node i; $Q$ is the consumption flow in leak situation.

$$RV(t) = \begin{bmatrix} p_1(t) - p_{1,0}(t) \\ \vdots \\ p_{ns}(t) - p_{ns,0}(t) \end{bmatrix} \tag{7}$$

where $p_i$ is observed pressure at sensor node i; $p_{i,0}$ is corrected model pressure in non-leak situation at sensor node i; $ns$ is the number of inner pressure sensors.

Second, fault signature matrix $FSM$ is built up, indicating the pressure responses to the leak in the model network. A leak (of specific flow) is modeled as extra demand added one by one to each node, the pressure drop at each sensor node is stored in the column of $FSM$.

$$FSM(t) = \begin{bmatrix} p_{1,1}(t) - p_{1,0}(t) & \cdots & p_{1,nn}(t) - p_{1,0}(t) \\ \vdots & \ddots & \vdots \\ p_{ns,nn}(t) - p_{ns,0}(t) & \cdots & p_{ns,nn}(t) - p_{ns,0}(t) \end{bmatrix} \tag{8}$$

where $nn$ is the number of nodes.

After that, each column of $FSM$ is correlated with $RV$ by Pearson's correlation coefficient to obtain the correlation vector rho, which is used to predict the leaky node.

$$\rho_{RV,FSM}(t) = \begin{bmatrix} \rho_{RV,FSM_1}(t) \\ \vdots \\ \rho_{RV,FSM_n}(t) \end{bmatrix} \tag{9}$$

$$\rho_{RV,FSM_j} = \frac{cov(RV, FSM_j)}{\sigma_{RV}\sigma_{FSM_j}} \tag{10}$$

where $\rho_{RV,FSM_j}$ is the correlation value of $RV$ and the column $j$ of $FSM$.

Eventually, all nodes are sorted by correlation values. The most probably leak node has the maximum correlation value.

$$\max_j \left( \rho_{RV,FSM_j}(t) \right), \ \ j = 1, \dots, n \tag{11}$$

The most likely leaky node can be predicted either at each scenario analysis time step $T_a$, or as the cumulative result during the entire leak event $T_w$.

Over the leak period $T_w$, for each node, the positive correlation values calculated at each time step $T_a$ are accumulated and averaged. By ranking the overall correlation values in descending order, the first node with the highest correlation value is the most probable leak location.

## 2.2 Graph-based interpolation

### 2.2.1 Theoretical background

Graph-based interpolation developed by X. Zhou et al. (2022) is an accurate and convenient approach for head estimation using limited pressure sensors in WDNs. The theoretical background of this method is Graph Signal Processing (GSP), using weights to represent the similarity between signals on adjacent nodes. GSP has a wide range of applications, such as sensor networks, image and 3D point cloud processing, and machine learning (Ortega et al., 2018). In this study, the structure of WDN is natural graph, including vertices and weighted edges corresponding to nodes $\mathcal{V} = \{1, 2, \dots, N\}$ and pipes $\mathcal{E} = \{w_{ij}\}_{i,j \in \mathcal{V}}$, respectively. Pressure head at each node

equipped with sensors are graph signals. Graph-based interpolation can infer unknown head values by not only considering the relation of the distance between observed data, but also hydraulic features.

In WDNs, if nodes i and j are connected then $w_{ij} > 0$, otherwise $w_{ij} = 0$. The adjacent matrix $\mathcal{W}$ of the weighted graph can be written as:

$$\mathcal{W}_{ij} = \begin{cases} w_{ij}, & \text{if } w_{ij} \in \mathcal{E} \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

The weight $w_{ij}$ quantifies the similarity of head signals between different nodes, which is assigned based on hydraulic features. X.Zhou et al. (2022) proposed four different ways to calculate weights, including uniform weight, nodal demand weight, Hazen-Williams weight, and model-simulated head loss weight. In this research, we use pipe length weight which requires less information from the model as described in Alves et al. (2022). Hazen-Williams weight is replaced with Darcy-Weisbach weight since Darcy-Weisbach formula is used in this study to calculate the head loss. Four weights are increasingly dependent on the network parameters.

Uniform weight $w_u$: $w_{ij} = 1$ (13)

Pipe length weight $w_l$: $w_{ij} = \frac{1}{l_{ij}}$, $l$ is the pipe length between node i and j (14)

Darcy-Weisbach weight $w_D$: $w_{ij} = \frac{l_{ij}}{D_{ij}}$, $D$ is the pipe diameter (15)

Head loss weight $w_h$: $w_{ij} = w_{ji} = \begin{cases} \frac{1}{h_{ij}}, h_{ij} > \varepsilon \\ \frac{1}{\varepsilon}, h_{ij} < \varepsilon \end{cases}$, $h_{ij}$ is the head loss of pipe ij calculated by model simulation and $\varepsilon$ is set at 0.001 to prevent from extremely big value. (16)

Head signals Head signals $H = [H_1, H_2, \dots, H_N]^T$ in WDN are considered original graph signals, which are decomposed onto a set of Laplacian spectrums $\mathcal{S} = [s_1, s_2, \dots, s_N]$ and corresponding eigenvectors $U = [u_1, u_2, \dots, u_N]$ by Graph Fourier Transform (GFT). In general, the variation of head signals at neighboring nodes is limited in WDNs, and it was proved that these 'smooth' changes of heads can be well explained by lower frequency and corresponding eigenvectors (Stanković et al., 2019; X. Zhou et al., 2022a). Smoothness of graph signals means higher similarity of head values (Ortega et al., 2018). Therefore, frequency limit $f$ is introduced to determine the low frequency part of $U$ and $H$ for approximation of the head signals over the entire WDN. Measured heads $ys$ can be considered as the original signal $H$. A matrix $Ds$ is created as a $M \times N$ sampling matrix with $Ds[i,j] = 1$ if the i[th] pressure sensor is installed at node j and 0 otherwise. The estimated low frequency spectrum is calculated as:

$$\mathcal{S}_f = pinv(DsU_f)ys, \qquad f \le M \tag{17}$$

The reconstructed heads are:

$$\hat{H} = U_f \mathcal{S}_f \tag{18}$$

In the study of X. Zhou et al. (2022), the interpolation was validated in real networks using synthetic data generated by EPANET simulator. Table 1 shows that the head loss weight performs the best among all weights with the mean absolute error 0.25 m in a middle-sized WDNs including 480 nodes and 567 pipes with 30 pressure sensors.

**Table 1.** Results of graph-based interpolation using different weights (X. Zhou et al. 2022).

| Graph weight | MAE [m] |
|---|---|
| Uniform | 0.509 |
| Demand | 0.289 |
| Hazen-W | 0.477 |
| Head loss | 0.254 |

### 2.2.2. Integration with model-based method

Based on the research in Quevedo et al. (2011), more pressure sensors improve the reliability (higher maximum correlation value) and accuracy of localization (as shown in Table 2). However, in practical cases, water utilities usually install a limited number of sensors due to their tight budgets. In this research, virtual sensors are created over the entire WDN by graph-based interpolation. After the interpolation, the number of sensors (ns) equals the number of nodes ($nn$) in the WDN. The size of RV and FSM will be enlarged from ns rows to $nn$ rows. Considering the accuracy of interpolation, the head loss weight is used to obtain the 'virtual' measurements of virtual sensors.

**Table 2.** Results of model-based method using different number of sensors (Quevedo et al. 2011).

| Number of sensors [ns] | Leak flow [l/s] | Noise level [m] | Max. correlation c [-] | Distance [m] |
|---|---|---|---|---|
| 3 | 6.3 | 0.25 | 0.460 | 224 |
| 15 | 6.3 | 0.25 | 0.546 | 109 |
| 100 | 6.3 | 0.25 | 0.686 | 68.7 |

## 2.3 Data-driven method with graph-based interpolation

In this work, a data-driven method is developed to locate the leak in DMAs with flow and pressure measurement at the inlet and a few pressure sensors inside. It is assumed that an efficient method has been applied to detect the leak. The method is modified from the research of Soldevila et al. (2019) and Soldevila et al. (2021), in which the authors compare the pressure residuals over the entire WDN in non-leak and leak situation. Graph-based interpolation is introduced to estimate the pressure at the nodes without sensors. Bayesian theory is used eventually for temporal reasoning to improve the localization results. In this case, a large dataset in the leak situation is not required, nor is an estimation of leak flow and a well-calibrated model, all of which are strong benefits. Figure 10 shows the overview of this method. The performance of the method is assessed two metrics described in section 2.4.

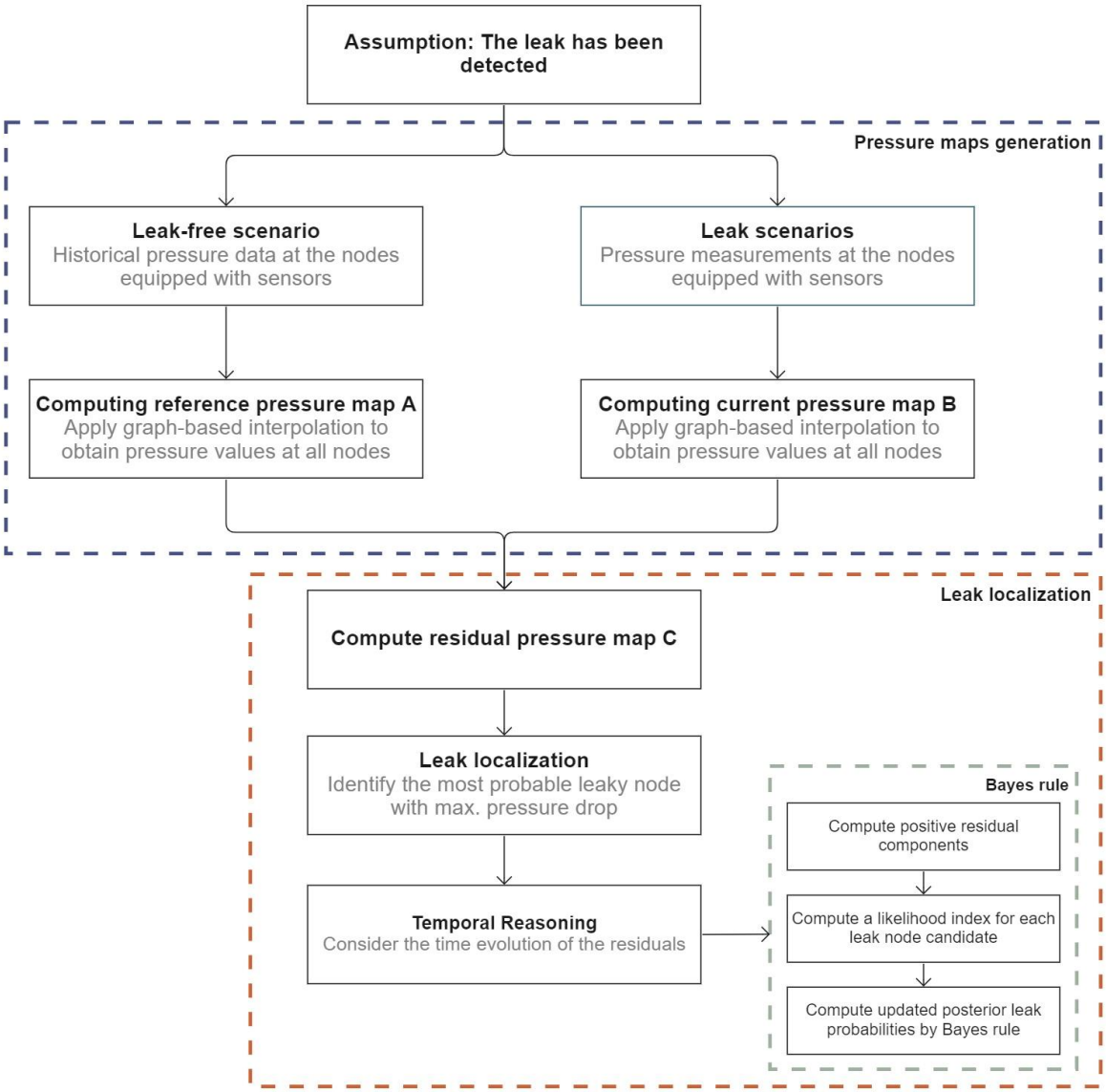


Figure 10. Overview of data-driven leak localization method.

Considering a WDN including $n$ nodes and $ns$ sensors, a residual vector $R = [R_1, R_2, \dots, R_n]^T$ is created at each time step $t$:

$$R_i(t) = p_i(t) - p_{li}(t) \tag{19}$$

where $p = [p_1, p_2, \dots, p_n]$ is the pressure map A in non-leak situation; $p_l = [p_{l1}, p_{l2}, \dots, p_{ln}]$ is the pressure map B in leak situation.

The leak location is identified by the maximum pressure residual (Romano et al., 2017a; Soldevila et al., 2019a)

$$J = \underset{i \in \{1, \dots, n\}}{argmax\{R(t)\}} \tag{20}$$

where $R_i$ is the item in vector $R$.

Soldevila et al. (2019) implemented this localization approach in the benchmark network Hanoi in the model environment. It is assumed that each node in the DMA is equipped with a sensor in an ideal situation. During the leak, the pressure residual profile is shown as Figure 11.

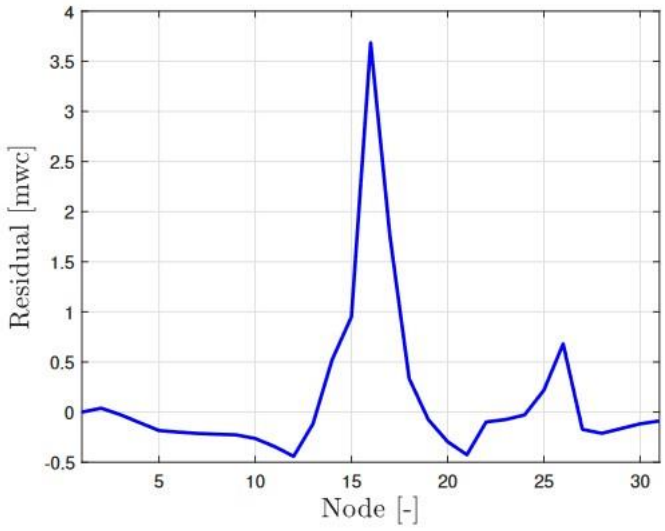


**Figure 11.** Pressure residual profile of the leak node (Soldevila et al., 2019).

In this study, graph-based interpolation is employed to obtain the pressure map in both non-leak and leak situations instead of the Kriging interpolation used by Soldevila et al. (2019) and Soldevila et al. (2021). It was proved that graph-based interpolation provides better head estimation than Kriging in X. Zhou et al. (2022) using fully synthetic data. Four different weights will be used to further explore the performance of graph-based interpolation with real measured data in WDNs.

The residual vector $R$ is computed using measured data at each time step. In this case, historical data from the same time period as the leak in the previous three days are considered as valid measured data as to obtain the pressure map in non-leak situation as the reference. In order to consider the time evolution of the residuals, Bayesian reasoning is introduced to improve the localization precision.

To avoid negative values, at each time step, the minimum component of $R(t)$ is added to every component i as an offset.

$$R_i^+(t) = R_i(t) + \min\left(R(t)\right) \tag{21}$$

Then a likelihood index $a_i(t)$ is computed for each node and be normalized as

$$a_i(t) = \frac{R_i^+(t)}{\sum_{j=1}^{n} R_j^+(t)} \tag{22}$$

and later be combined with the initial leak probabilities for each node through Bayes rule. The updated posterior probability of leak is

$$P_i(t) = \frac{P_i(k-1)a_i(t)}{\sum_{j=1}^{n} P_j(k-1)a_j(t)} \tag{23}$$

The initial probability of leak is equal for every node as $P_i(k) = 1/n$. The leaky node has the highest probability of leak.

$$J(t) = \underset{i \in \{1,\dots,n\}}{argmax}\{P_i(k)\} \tag{24}$$

## 2.4 Performance metrics

Compared with leak detection, there are not many performance metrics proposed to evaluate the localization accuracy. Generally, most researchers use either the distance or the pipe length between the actual leak and the predicted leak. Moors et al. (2018) proposed false-positive rate to evaluate the localization results. It is the percentage of false positive nodes, which are the nodes with higher correlation value than the real leak node. This metric can show to what extent the localization approach can help reduce the search area of the leak.

In this study, two novel metrics are proposed to assess the performance of the methodology: localization error (LE)and false positive path (FP path). LE focuses on nodes in WDNs. We draw a circle on the map with the predicted leak node as the center and the distance between actual and predicted leak node as the radius (Figure 12). LE is the percentage of nodes inside the circle to all nodes. It should be as low as possible, indicating smaller search area of the leak and shorter linear distance. LE is used to evaluate the performance of both model-based and data-driven methods.

$$LE = \frac{Nodes\ inside\ the\ circle\ (center = max.corr\ node; r = distance\ to\ real\ leak)}{Total\ nodes\ in\ the\ DMA} \times 100\% \tag{25}$$

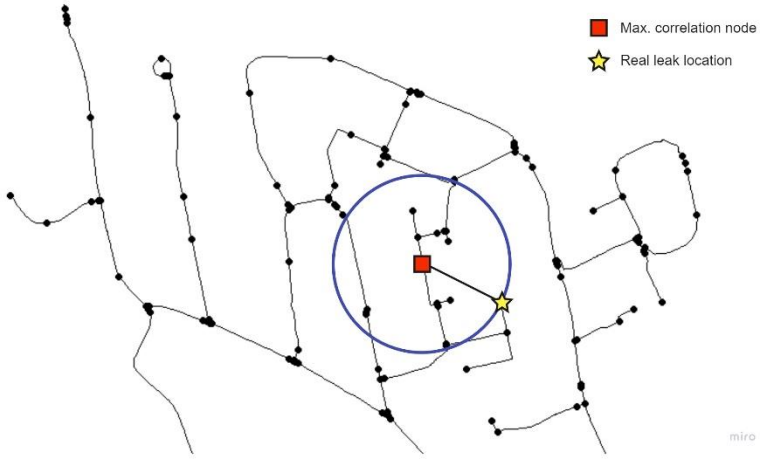


**Figure 12.** Conception drawing of localization error.

However, nodes inside the circle maybe close in space but far away from each other from a hydraulic point of view. In addition, the uneven distribution of nodes in WDNs also affects the fairness of this metrics. Consequently, in more central areas, where nodes are more concentrated, LE is probably higher than in remote areas. But the distance between the leak and the predicted leak node is the same.

FP path is calculated based on false positive nodes and pipe length in WDNs. In practice, when we search for the leak using model-based method, the node with the highest correlation value should be looked at first, and then

the second highest, the third, etc., until we find the actual leak node. The group of nodes that have higher correlation values than the actual leak node are false positive nodes. All pipes connected to those false positive nodes are possible search pipes. To avoid double counting one pipe from two nodes, only half of the pipe length is counted for one node. FP path equals the percentage of search pipe length to total pipe length, which should be as small as possible.

$$FP\ path = \frac{Possible\ search\ pipe\ length\ [m]}{Total\ pipe\ length\ in\ DMA\ \ [m]} \times 100\% \tag{26}$$

where possible search pipe length is the sum of half of all pipes connected to false positive nodes.

A possible drawback is the sparse distribution of false positive nodes. They are not always close to and around the leak and a few maybe far away from the leak node. In this case, the actual search area is larger than the FP path shows.

The results of both LE and FP path are split into four classes to improve the fairness of assessment. If LE or FP path < 10%, it is interpreted as good, 10 – 20% as fair, between 20 - 40% as acceptable, and > 40% as poor.

For data-driven method, the hydraulic pipe length between actual leak and predicted leak node is also used for performance evaluation.

# Chapter 3: Case Study

SES water is a water company in the UK, providing drinking water services to an area of south London divided into over 300 DMAs with more than 750,000 people as shown in Figure 13. SES Water aims to reduce NRW by 15% from 2020 to 2025. They plan to detect and find leaks sooner by using software-based solutions since they are unable to renew infrastructure very often with a tight budget. 10 of DMAs are selected to test and validate the two approaches for leak localization. Each one of them is supplied by one inlet and equipped with one flow meter and a pressure sensor at the inlet and a pressure sensor at the critical point as well as additional 5 - 8 inner pressure sensors.

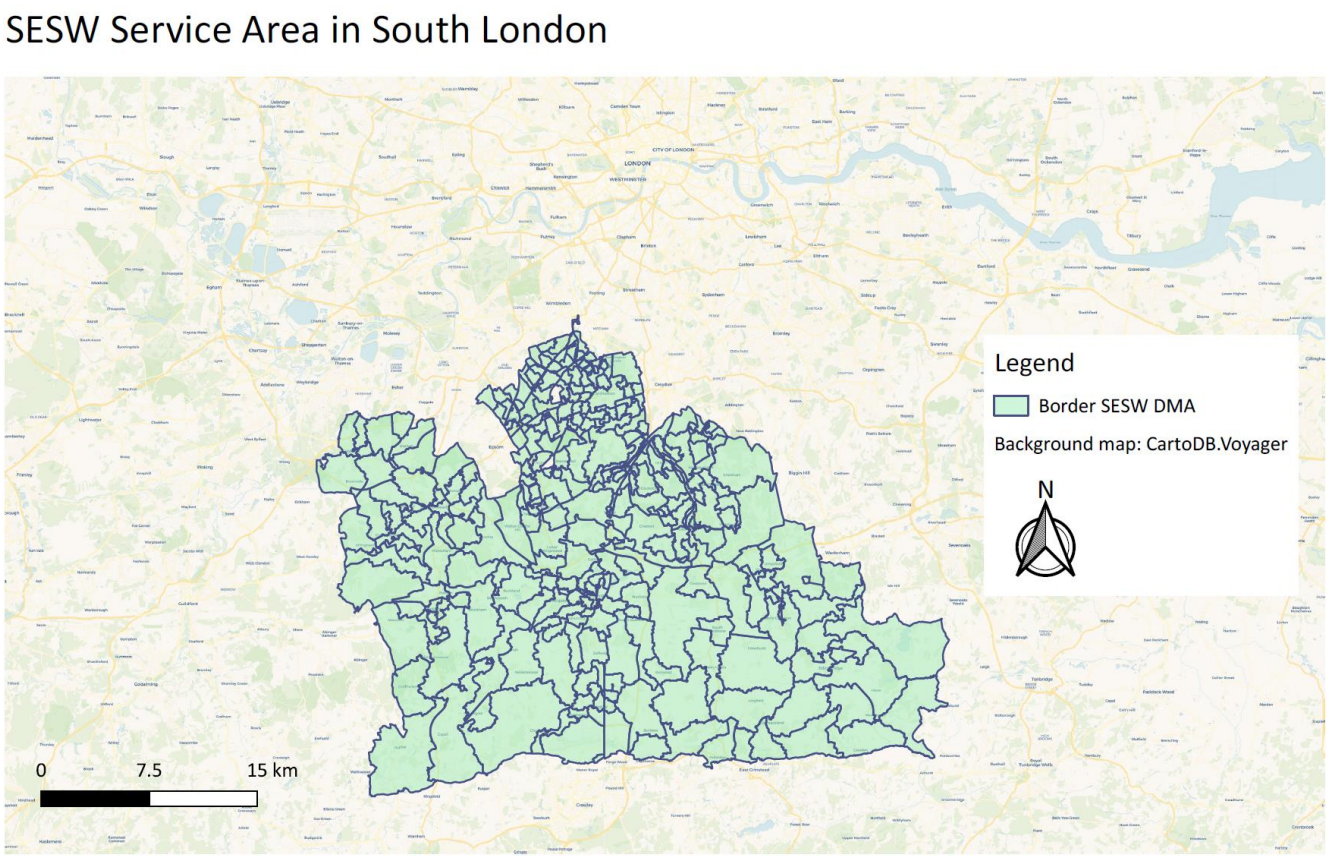


**Figure 13.** Case study map: SES Water's service area in South London, UK.

## 3.1 Description of studied DMAs

Table 3 shows the basic information of these 10 DMAs.

**Table 3.** Overview of 10 DMAs in South London, UK.

| DMA | Average flow [l/s] | Sensors | Land use | Elevation difference [m] | Node | Pipe length [km] |
|---|---|---|---|---|---|---|
| DMA1 | 1.36 | 6 | Rural | 45.6 | 88 | 7.2 |
| DMA2 | 1.96 | 6 | Rural | 24.4 | 138 | 10.8 |
| DMA3 | 2.40 | 6 | Urban | 12.6 | 96 | 3.9 |
| DMA4 | 3.76 | 7 | Urban | 3.0 | 114 | 3.5 |
| DMA5 | 3.76 | 7 | Semi-Urban | 38.1 | 215 | 11.9 |
| DMA6 | 4.27 | 7 | Urban | 34.6 | 127 | 5.5 |
| DMA7 | 5.50 | 8 | Urban | 46.8 | 285 | 10.2 |
| DMA8 | 7.02 | 8 | Semi-Urban | 46.9 | 343 | 24.9 |
| DMA9 | 9.15 | 8 | Semi-Urban | 15.7 | 314 | 13.0 |
| DMA10 | 11.68 | 9 | Urban | 16.4 | 337 | 11.6 |

*Sensors: The number of pressure sensors in each DMA; Elevation: The elevation difference between the lowest and the highest point in each DMA.

All DMAs cover the total pipe length of 121.6 km and 2555 nodes. The selected DMAs have varying sizes from 88 to 337 user nodes and different characteristics including the type of WDN, land use, and topography. A total of 87 pressure sensors are placed with an average of 7 sensors in each DMA, respectively. Each pressure sensor covers 13 to 34 nodes. The average flow of DMA varies greatly from 1.36 l/s (DMA1) to 11.68 l/s (DMA10). Most DMAs are hybrids with branch and loop structures as shown in Figure 14, which is very different from the benchmark networks usually applied for test and validation of leak localization strategies. DMA1 and DMA2 are located in rural areas with farms around, and others are typically urban areas. The topographic elevation difference ranges from 3 m to 62 m. 7 DMAs have an elevation difference of more than 30 m, where the pressure has considerable validation.

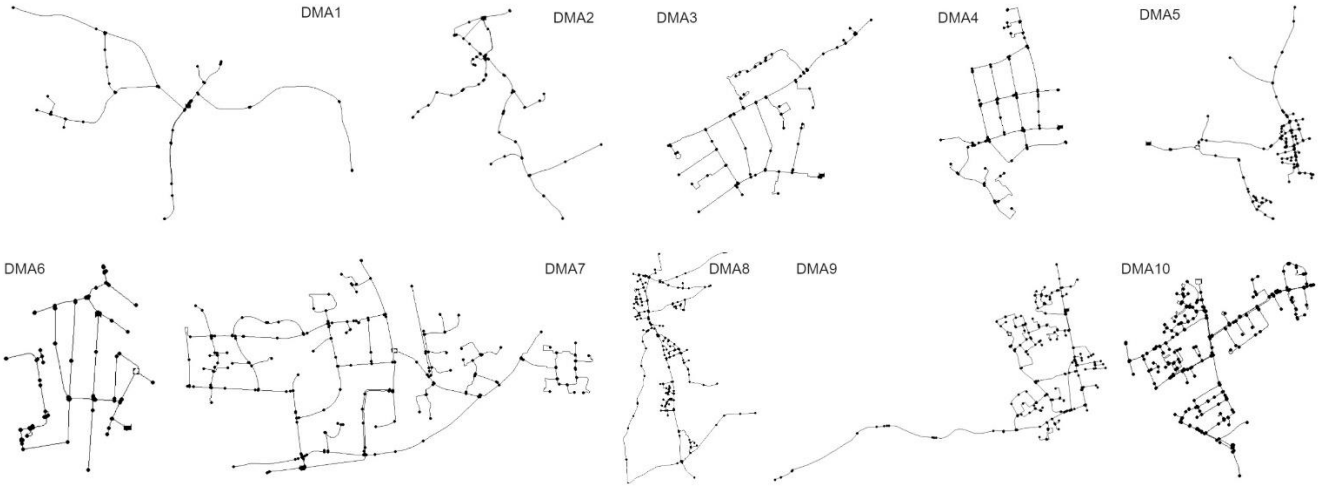


**Figure 14.** Network maps of 10 DMAs.

## 3.2 Leak tests plan

Artificial leak tests are performed by technicians from SES Water by opening fire hydrants.

A total of 22 leak tests including 47 leak events were performed in two stages as follows:

- A leak test was conducted from Feb 03 to Apr 01 in each of the 10 DMAs at randomly selected locations. The protocol of the leak test is the same in leak size and duration but different in the execution time. Different sizes of leaks were controlled by how much the fire hydrant is opened. Specifically, constant leak flow leaks of 0.8 or 1.5 l/s (i.e., most between 20% and 40% of the average inflow of the DMA) with durations of either 1.5 or 1 hours were created, respectively. Leak tests were done either in the morning from 9:00 to 11:30 or from 13:00 to 15:30.
- 15 extra leak tests were carried out at 7 DMAs from May 07 to June 01 in order to further investigate different impact factors, including the longer duration, higher leak flow, nighttime test, and leak location (either in loop or branch area). Nighttime tests with the same leak flow and duration were performed in two DMAs. Different leak locations were explored in two larger DMAs with hybrid network structures. Higher leak flows of 2.0 and 2.5 l/s, as well as a longer duration (6 hours) of small leak of 0.8 l/s, were also scheduled.

The leak tests were performed by ensuring clients were subject to as few disruptions as possible in the process. To achieve this, leak tests were scheduled to avoid the morning and evening peak hours to make sure the water supply service. The maximum test flow allowed to be performed was 2.5 l/s in the daytime. Table 4 and Table 5 shows the plan of leak tests in detailed.

**Table 4.** Overview of leak test scenarios in the first stage.

| Date | DMA | Time | Test flow [l/s] | DMA | Time | Test flow [l/s] |
|---|---|---|---|---|---|---|
| 3-2-2022 | DMA2 | 09:00-10:30 | 0.8 | DMA5 | 13:00-14:30 | 0.8 |
| | | 10:30-11:30 | 1.5 | | 14:30-15:30 | 1.5 |
| 10-3-2022 | DMA8 | 09:00-10:30 | 0.8 | DMA1 | 13:00-14:30 | 0.8 |
| | | 10:30-11:30 | 1.5 | | 14:30-15:30 | 1.5 |
| 15-3-2022 | DMA6 | 09:00-10:30 | 0.8 | DMA3 | 13:00-14:30 | 0.8 |
| | | 10:30-11:30 | 1.5 | | 14:30-15:30 | 1.5 |
| 23-3-2022 | DMA10 | 09:00-10:30 | 0.8 | DMA4 | 13:00-14:30 | 0.8 |
| | | 10:30-11:30 | 1.5 | | 14:30-15:30 | 1.5 |
| 24-3-2022 | DMA7 | 09:00-10:30 | 0.8 | DMA9 | 09:00-10:30 | 0.8 |
| | | 10:30-11:30 | 1.5 | | 10:30-11:30 | 1.5 |

**Table 5.** Overview of leak test scenarios in the second stage.

| Date | DMA | Time | Test flow [l/s] | Research goal |
|---|---|---|---|---|
| 5-5-2022 | DMA3 | 09:00-10:30 | 0.8 | Different times of the leak.<br>a. Nighttime tests in DMA3 and DMA6<br>b. Switching time leak tests in DMA3, DMA6, DMA1, DMA9, and DMA5. |
| | | 10:30-11:30 | 1.5 | |
| 1-6-2022 | DMA3 | 01:00-02:30 | 0.8 | |
| | | 02:30-03:30 | 1.5 | |
| 3-5-2022 | DMA6 | 09:00-10:30 | 0.8 | |
| | | 10:30-11:30 | 1.5 | |
| 4-5-2022 | DMA6 | 01:00-02:30 | 0.8 | |
| | | 02:30-03:30 | 1.5 | |

| | | | | |
|---|---|---|---|---|
| 10-5-2022 | DMA1 | 09:00-10:30 | 0.8 | |
| | | 10:30-11:30 | 1.5 | |
| 12-5-2022 | DMA9 | 13:00-14:30 | 0.8 | |
| | | 14:30-15:30 | 1.5 | |
| 13-5-2022 | DMA5 | 09:00-10:30 | 0.8 | |
| | | 10:30-11:30 | 1.5 | |
| 16-5-2022 | DMA10 | 09:00-10:30 | 0.8 | Different leak locations in hybrid networks.<br>a. Due to the larger service area of DMAs, DMA10 and DMA7, as well as the more complex layout of the networks, these two DMAs are chosen for additional leak tests |
| | | 10:30-11:30 | 1.5 | |
| 17-5-2022 | DMA10 | 09:00-10:30 | 0.8 | |
| | | 10:30-11:30 | 1.5 | |
| 26-5-2022 | DMA7 | 09:00-10:30 | 0.8 | |
| | | 10:30-11:30 | 1.5 | |
| 27-5-2022 | DMA7 | 09:00-10:30 | 0.8 | |
| | | 10:30-11:30 | 1.5 | |
| 24-5-2022 | DMA9 | 09:00-15:00 | 0.8 | Longer duration of small leak flow. |
| 25-5-2022 | DMA4 | 09:00-15:00 | 0.8 | |
| 18-5-2022 | DMA3 | 11:00-12:00 | 0.8 | Bigger leak size. |
| | | 12:00-13:00 | 1.5 | |
| | | 13:00-14:00 | 2.0 | |
| | | 14:00-15:00 | 2.5 | |
| 23-5-2022 | DMA4 | 11:00-12:00 | 0.8 | |
| | | 12:00-13:00 | 1.5 | |
| | | 13:00-14:00 | 2.0 | |
| | | 14:00-15:00 | 2.5 | |

# Chapter 4: Results and Discussions

## 4.1 Sensor evaluation

Before the localization analysis, every sensor should be assessed based on MAPE in Eq. (5) before and after model output correction in the non-leak situation (see section 2.1.4). An overview of sensor evaluation in DMA3 using 72-h data in the non-leak situation is shown as an example in Table 6. In Figure 15, the observed and model-simulated head at sensor P1 node were obtained from Mar 09 to Mar 12 (3 weekdays) in 2022. Ideally, observed heads are aligned with model simulated values shown as the best-fitting line (red) in Figure 16 (Left). In practice, data points deviate slightly from the best-fitting line in the presence of uncertainties. A quadratic residual model fitted to 72-h flow and residuals is used to correct model outputs. However, residuals still exist and the MAPE1 after calibration is still higher than 0.25%. Therefore, sensor P1 is considered a faulty sensor and thus is not used in localization analysis. The other five sensors are valid sensors with MAPE1 < 0.25%.

**Table 6.** Overview of the sensor evaluation in DMA3

| Name | MAPE [%] | C1 | C2 | MAPE1 [%] | Remove sensor? [Yes/No] |
|---|---|---|---|---|---|
| DMA3_P1 | 0.94 | -118340 | -0.01 | 0.35 | Yes |
| DMA3_P2 | 1.27 | -103481 | -0.39 | 0.12 | No |
| DMA3_P3 | 0.67 | -100059 | 0.09 | 0.09 | No |
| DMA3_P4 | 0.75 | -101122 | 0.03 | 0.09 | No |
| DMA3_P5 | 0.37 | -101465 | 0.35 | 0.09 | No |
| DMA3_CP | 1.84 | -97233 | 2.21 | 0.10 | No |

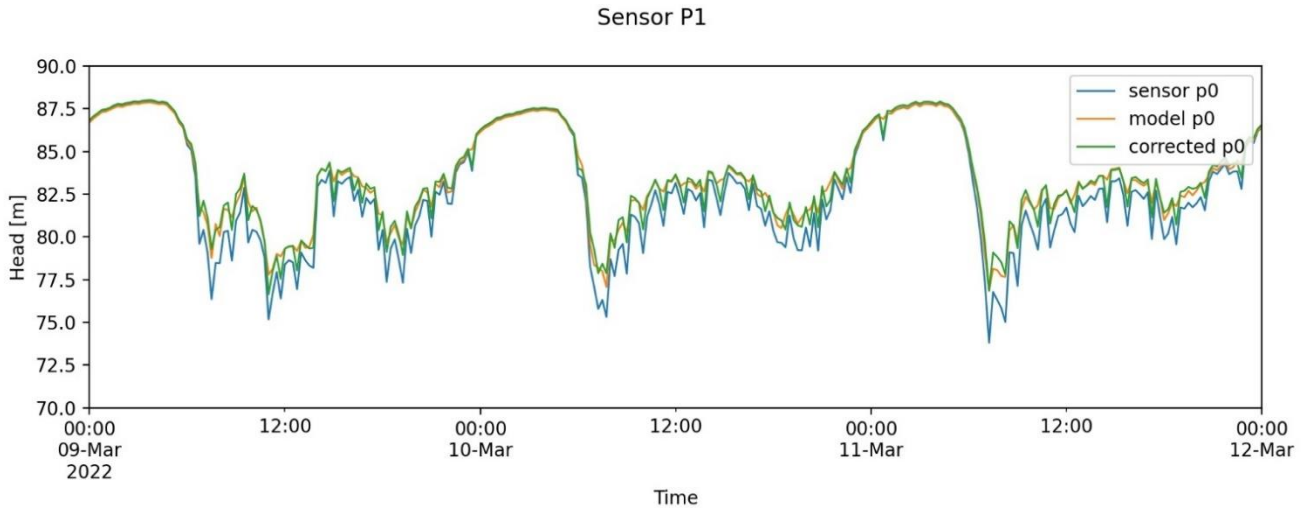


**Figure 15.** The time evolution of pressure head at sensor P1 node in DMA3. Blue and orange lines represent the observed and model-simulated data, respectively. The green line stands for model output after correction.

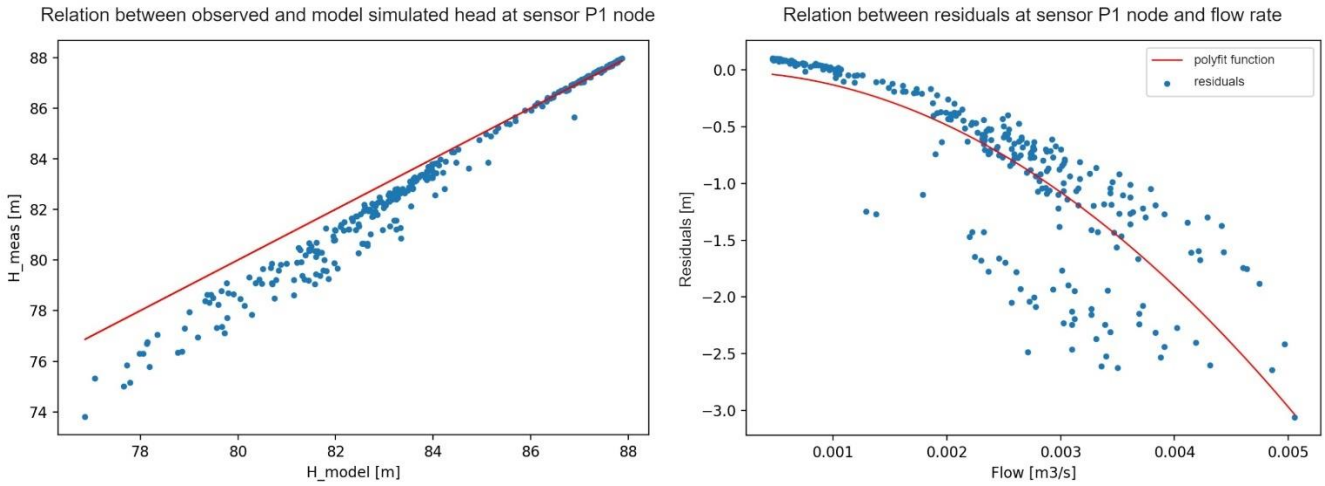


**Figure 16.** Left: Model-simulated and observed heads (sum of measured pressure and elevation of the node) in the non-leak situation; The red line represents the ideal situation with an accurate model and perfect sensor. Right: The relation between residuals and flow rate.

In total, there are 72 pressure sensors installed in 10 DMAs. Figure 17 illustrates the MAPE of each sensor in the non-leak situation before and after model output correction. The average MAPE of all sensors is 1.37%, in which case pressure changes driven by small leaks can be easily masked by residuals leading to poor results (Sophocleous et al., 2019). Before model output correction, only 15 sensors have MAPE < 0.25%. The MAPE of sensor P1 in DMA4 is even higher than 20%. As a result, we can see network models in these case studies cannot always represent the hydraulic status in real life due to uncertainties, particularly in nodal demand. The average MAPE1 is 0.19% after model output correction, which is reduced by 86.6%. Based on the sensor evaluation results, 13 sensors that show abnormal behavior are considered faulty sensors. One faulty sensor in each DMA on average. Consequently, they are discarded before localization analysis. Detailed information on sensor evaluation is shown in Appendix 1.

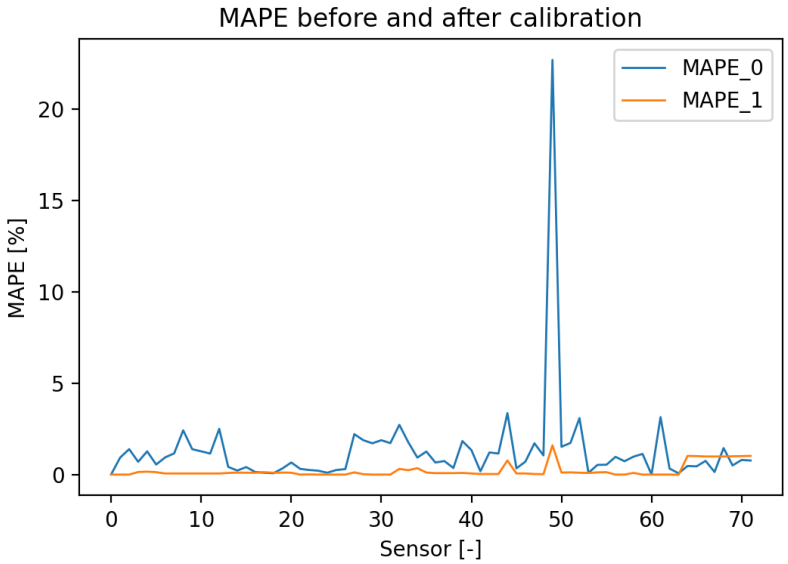


**Figure 17.** MAPE of each sensor in non-leak situation before (blue) and after (orange) model output correction.

In DMA9, there are 8 pressure sensors distributed in the network. However, the observed values at each sensor node show very different trends and shapes from model outputs as reported in Figure 18, especially during daytime. Therefore, the network model of DMA9 cannot represent the real situation in WDN. The deviations in the model simulated and observed values cannot be easily modeled as quadratic function related to flow. It can probably be explained by the wrong sensor settings. Consequently, DMA9 is considered a biased DMA equipped with faulty sensors. Leak tests performed in this DMA are considered invalid and are not included in the subsequent analysis of localization results.

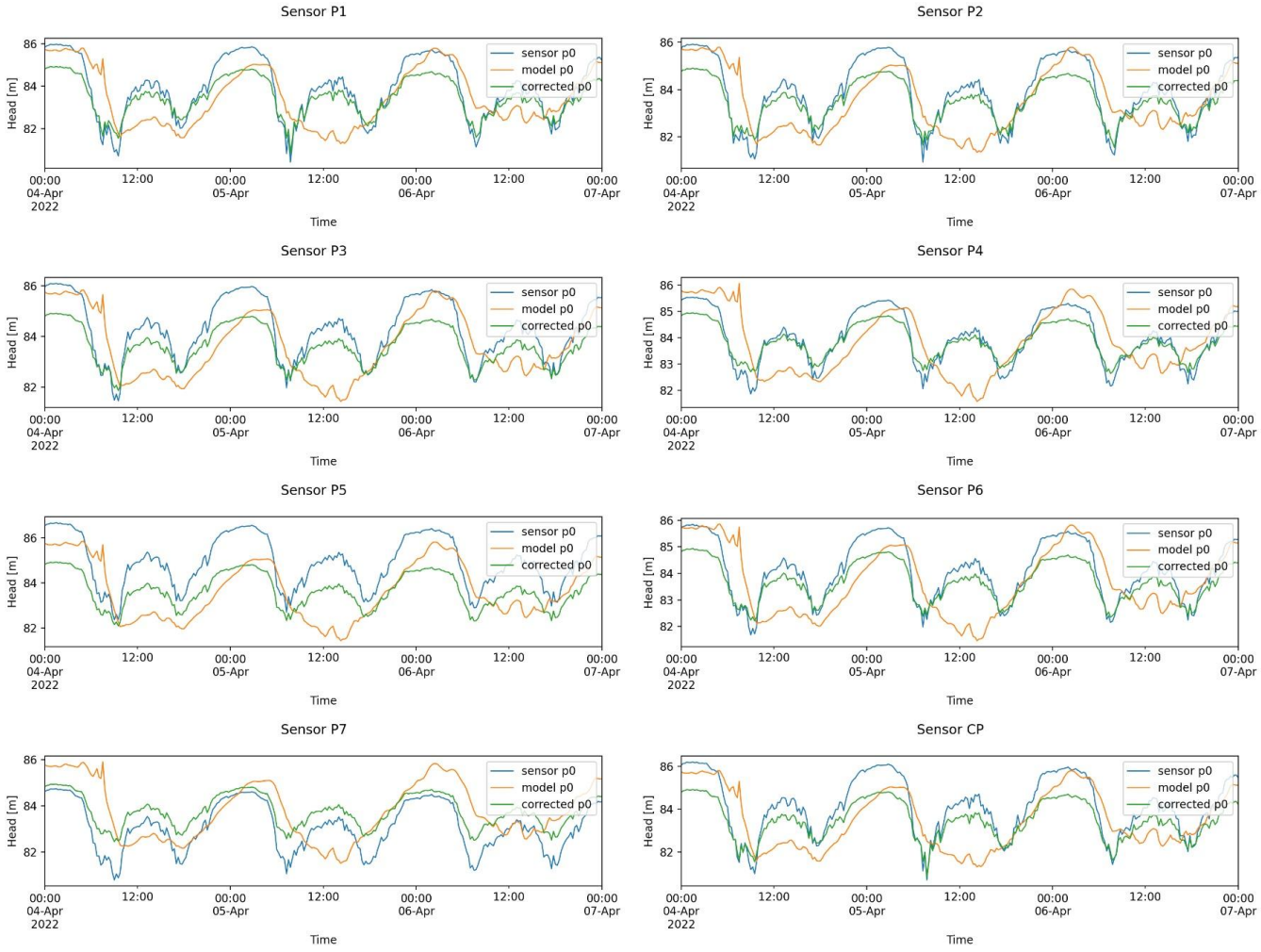


**Figure 18.** The observed pressure head (blue), model simulated data (orange), and the corrected values (green) in the non-leak situation in DMA9.

## 4.2 Graph-based interpolation

In this study, four different graph weights are used for graph-based interpolation (defined in section 2.2.1) to obtain the pressure heads at the nodes without a sensor. The performance of interpolation is evaluated by mean absolute error (MAE) as:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\overline{H_i} - H_i|$$

where $\overline{H_i}$ is the interpolated head of node, $N$ is the number of unknown nodes and $H_i$ is the model simulated head at the node without sensor. In this interpolation process, the known head signal at each sensor node is generated by model simulation in EPANET using boundary flow and pressure in the non-leak situation. The connectivity and other model parameters such as pipe diameter and pipe length are provided from the SES Water. Head loss calculated by the EPANET simulator at the morning peak hour (7 am) is used to determine the head loss weight of each pipe.

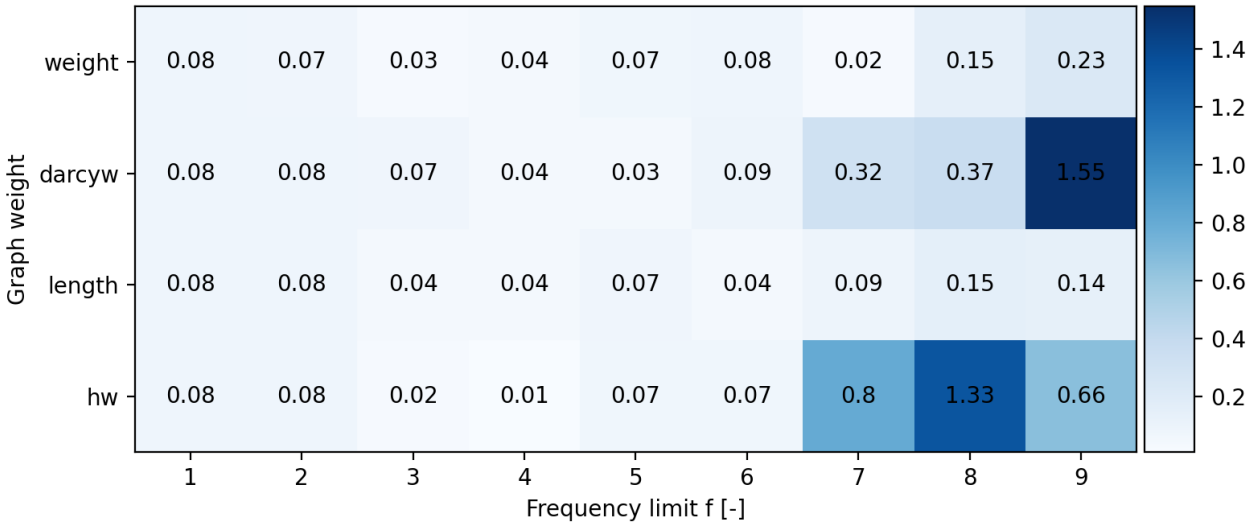


**Figure 19.** MAE under different combinations of graph weight w and frequency limit $f$. Unit: m; Deeper blue stands for higher MAE.

Figure 19 shows interpolation results in DMA7 using different combinations of graph weight $w$ and frequency limit $f$ in non-leak situation. DMA8 contains 285 nodes, and a total pipe length of 10.2 km. 8 pressure sensors are installed in the DMA. The largest head loss in DMA7 is only 0.17 m, which indicates the slow varying and very smooth change of head signals among different nodes. In the middled-size DMA described in Zhou et al., (2022), there are 480 nodes with 30 pressure meters installed. In comparison, the low spatial density of pressure sensors in DMA7 can be reflected by the low ratio of sensors to nodes, which is an unfavorable condition for interpolation.

The MAE obtained by graph-based interpolation in DMA7 range from 0.01 to 1.55 m. In general, higher MAE corresponds to a larger frequency limit $f$. According to the research by Zhou et al., (2022), graph-based interpolation performs better with slow varying heads as they are more aligned with the smoothness hypothesis of this method. Larger $f$ can introduce fast varying basic vectors, resulting in the decrease of stability of the

interpolation results. However, if the $f$ is too small, head changes in the WDN cannot be well translated into varying basic vectors and thus accurate interpolation cannot be obtained. For DMA7, the lowest MAE of 0.01 m is achieved when using the head loss weight and a frequency limit of 4.

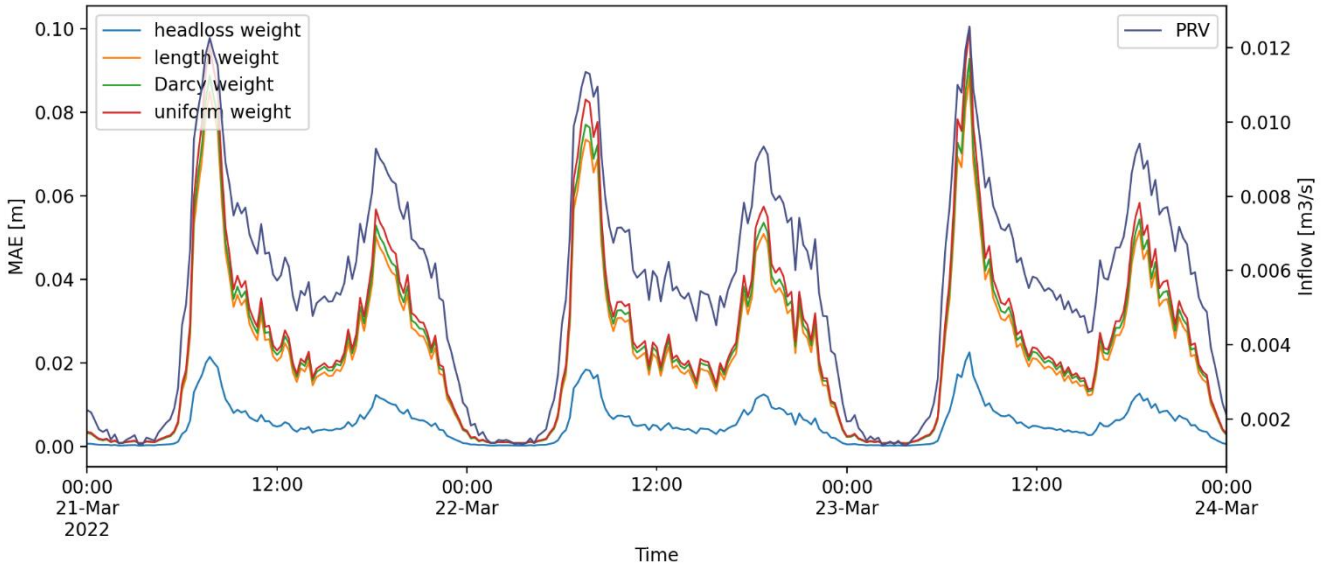


**Figure 20.** The MAE of interpolation using different graph weights at different times of three days in DMA7.

Figure 20 shows the variation of MAE in DMA7 over time with the optimal frequency limit $f$ at 4. In general, the pattern of MAE obtained for each graph weight changes the same as the flow pattern. The accuracy of interpolation is better with the smaller flow and goes lower with higher flow. The flow rate has less impact on the MAE of interpolation through the use of head loss weight compared with the other weights. The difference between the MAE obtained from the interpolation using the head loss weight and the MAE of the other weights is smallest at night, when the flow rate is the lowest during the day. Overall, the head loss weight outperforms the others due to the most accurate estimation of the relationship of heads at different nodes. In conclusion, the performance of interpolation can be greatly improved by adding more model information. Based on the equation for different weights, the uniform weight, length weight, and Darcy weight are increasingly accurate explanation of the hydraulic features among each node. However, there is not much difference between the MAE of uniform weight, length weight, and Darcy weight.

Table 7 shows the best selection of graph weight $w$ and frequency limit $f$ in 10 DMAs, in which the graph-based interpolation achieves the lowest MAE. Apart from DMA1, head loss weight performs the best in other DMAs, which is similar to the results of Zhou et al. (2022). In DMA2, it is not only length weight gives the best results with MAE=0.02 m, but other weights also perform similarly with MAE from 0.03 – 0.08 m. The optimal frequency limit $f$ in 10 DMAs varies from 33.3% to 83.3% of the number of sensors, and it mostly concentrated in the range of 40 - 60%. In Zhou et al., (2022), the authors recommend the frequency limit $f$ as 90% of the number of sensors, which is very different from the results in our case studies. The reason can be explained by the different magnitude of head loss in this research. The max. head loss in the case DMA is 6.12 m in the study of Zhou et al., (2022), and in the case of 10 DMAs the max. head loss ranges from 0.07 to 2.66 m. Therefore, lower frequency limit $f$ can better describe the varying of head signals in this study. Considering the measurement errors or inappropriate setting of sensor locations in real-life applications, the smaller f set as 50% of the number of sensors is recommended here.

Table 7. The best selection of graph weight and the frequency limit for the lowest MAE in 10 DMAs.

| DMA | Graph weight | Frequency limit f | Min. MAE [m] | Sensors | f / sensors [%] | Nodes |
|---|---|---|---|---|---|---|
| DMA1 | head loss | 5 | 0.00 | 6 | 83.3 | 88 |
| DMA2 | length | 3 | 0.02 | 6 | 50.0 | 138 |
| DMA3 | head loss | 3 | 0.01 | 6 | 50.0 | 96 |
| DMA4 | head loss | 4 | 0.01 | 7 | 57.1 | 114 |
| DMA5 | head loss | 3 | 0.03 | 7 | 42.9 | 215 |
| DMA6 | head loss | 3 | 0.09 | 7 | 42.9 | 127 |
| DMA7 | head loss | 4 | 0.01 | 8 | 50.0 | 285 |
| DMA8 | head loss | 5 | 0.05 | 8 | 50.0 | 343 |
| DMA9 | head loss | 3 | 0.04 | 8 | 37.5 | 314 |
| DMA10 | head loss | 3 | 0.07 | 9 | 33.3 | 337 |

* MAE is calculated using 72-h data in the non-leak situation; Sensor: Total number of sensors including the pressure sensor at the boundary of the DMA.

## 4.3 Model-based leak localization

### 4.3.1 General results of model-based method

The model-based method with/without model output correction is applied to 10 different DMAs in South London, UK. Five leak events in DMA9 are not included due to faulty pressure sensors (see section 4.1). Therefore, there are 47 leak events in total taken into account (see section 3). The duration of these artificial leak events is 1 hour or 1.5 hours or 6 hours. One-minute step pressure and flow data are provided by the water company. In the study of this section, the scenario analysis time step is 15 minutes, which means the original data at a resolution of one minute is resampled to 15 minutes. For example, four localization results are generated and averaged in an hour leak event to get an overall result.

The performance of model-based leak localization is reported in Table 8. Both metrics LE and FP path indicate the percentage of search area of the leak (see section 2.4). Table 8 shows the percentage of leak events corresponding to each of the four result classes under each strategy. BASE stands for no model output correction and faulty sensors removal; RE stands for removing faulty sensors; MP means model output correction.

**Table 8.** The localization results of the percentage of leak events in different result classes.

| Results category | BASE | | RE | | MP | | MP+RE | |
|---|---|---|---|---|---|---|---|---|
| | LE | FP path | LE | FP path | LE | FP path | LE | FP path |
| Good [0-10%] | 8.5% | 23.4% | 10.6% | 10.6% | 29.8% | 57.4% | 40.4% | 63.8% |
| Fair [10-20%] | 12.8% | 2.1% | 12.8% | 4.3% | 29.8% | 17.0% | 27.7% | 14.9% |
| Acceptable [20-40%] | 31.9% | 51.1% | 44.7% | 44.7% | 19.1% | 10.6% | 17.0% | 21.3% |
| Poor [>40%] | 46.8% | 23.4% | 31.9% | 40.4% | 21.3% | 14.9% | 14.9% | 0.0% |

Results in Table 8 illustrate that the model output correction and faulty sensor removal have a significant positive impact on the performance of model-based leak localization. With the help of model output correction and faulty sensor removal (MP+RE), for 68.1% of leak events, the final candidates involved only less than 20% of all nodes as possible leak locations from the view of localization error (LE). Considering the false positive path (FP path), for 78.7% of leak events, the leak is possible to be found within 20% of the total pipe length in the WDN. In contrast, localization results of more than 70% of leak events fall into 'Acceptable' or 'Poor' classes using the BASE strategy. The general improvement in performance is most likely explained by the reduction of differences between observed and model-simulated data by model output correction and faulty sensor removal.

The average pressure drop of all leak events is 0.55 m. However, according to findings in section 4.1, before model output correction, the mean error between observed and model-simulated data in the non-leak situation is 1.34 m transformed by a value of 1.37% of the average MAPE. In this case, it is difficult to identify the leak in which node causes the pressure drop because it can be masked by the difference due to the mismatch of model and reality (Sophocleous et al., 2019). The mean error is reduced to 0.16 m by model output correction, in which case the pressure response to the leak is more prominent, making it easier to find the leak. The percentage of leak events with poor results decreases greatly from 46.8% to 14.9% for LE and from 23.4% to 0.0% for FP path.

Table 9. Performance of four strategies in different DMAs.

| DMA | Leak events | Average flow [l/s] | BASE | | RE | | MP | | MP+RE | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | LE | FP path | LE | FP path | LE | FP path | LE | FP path |
| DMA1 | 4 | 1.4 | 55.8% | 23.5% | - | - | 1.0% | 2.3% | - | - |
| DMA2 | 2 | 2.0 | 9.0% | 96.0% | - | - | 41.0% | 32.0% | - | - |
| DMA3 | 10 | 2.4 | 37.6% | 26.8% | 44.5% | 34.4% | 7.4% | 9.1% | 5.8% | 5.6% |
| DMA4 | 7 | 3.8 | 53.7% | 42.4% | 26.9% | 64.0% | 36.7% | 18.7% | 17.9% | 14.9% |
| DMA5 | 4 | 3.8 | 32.0% | 7.5% | - | - | 30.0% | 24.0% | - | - |
| DMA6 | 6 | 4.3 | 18.5% | 2.1% | 19.7% | 53.0% | 21.7% | 2.3% | 40.0% | 13.7% |
| DMA7 | 6 | 5.5 | 35.3% | 49.0% | - | - | 13.3% | 7.8% | - | - |
| DMA8 | 2 | 7.0 | 60.5% | 30.0% | - | - | 18.0% | 1.0% | - | - |
| DMA10 | 6 | 11.7 | 75.0% | 56.3% | 54.5% | 29.2% | 71.5% | 52.0% | 43.0% | 12.7% |
| Average | | | 41.9% | 37.1% | 37.6% | 43.0% | 26.7% | 16.6% | 23.4% | 12.7% |

*'-'means no faulty sensor in the DMA; The average results of all leak events using RE or MP+RE also involve the results in the DMA without faulty sensors.

Table 9 shows the performance of four strategies in 9 DMAs, respectively. Both metrics LE and FP path are average results of all leak events in each DMA. Comparing the results using the strategy of BASE and MP, it can be seen that the overall performance of MP strategy is better than or similar to that of BASE strategy in 8 DMAs except for DMA2. For DMA2, although LE is only 9.0% using the BASE strategy, FP path is extremely high at 96.0%, which means the correlation value of the leak node is very low, and the final list of pipe candidates involved almost all pipes in the network. In this case, the predicted leak node with the maximum correlation value near the actual leak node is probably a spurious event. Results in the other 8 DMAs prove that model output correction can improve localization performance.

Six faulty sensors are distributed in DMA3, DMA4, DMA6, and DMA10, three of which are in DMA6 and one in each of the remaining three DMAs. Table 9 illustrates that only in DMA10 is the improvement in localization precision more obvious by removing faulty sensors (RE). The performance in DMA3 and DMA6 is even worse. The possible reason is the very limited number of pressure sensors leads to insufficient sensitivity to pressure changes in the system. Although some sensors are not accurate enough, they still help to capture the leak-driven pressure changes as much as possible. However, comparing the results using the strategy of MP and MP+RE, the performance of MP+RE is slightly better in DMA3 and greatly in DMA4 and DMA10. The poor results in DMA6 can be probably explained by the inappropriate removal of faulty sensors which are excluded based on the criteria defined in section 2.1.4. Therefore, only four valid sensors remain in DMA6, and the actual leak node is very close to one of the removed sensors P1 as shown in Figure 21.

These findings above give two insights into the application of the methodology. First, comparing the results of the four different strategies together, MP+RE still has the best performance in general. Compared to the average result of all leak events from the strategy BASE, MP+RE outperforms in terms of LE and FP path by 44.9% and 61.1%, respectively. The applicability of MP+RE in different types of DMAs in different conditions can be confirmed by the best results obtained in seven DMAs. Model output correction contributes more to improved localization performance than faulty sensor removal does. Second, in practice, the criteria for faulty sensors should be determined with great care. In this study, the criteria work well in most case studies, but it has not been tested in other WDNs such as in other countries or owned by other water companies, or equipped with other brands of sensors.

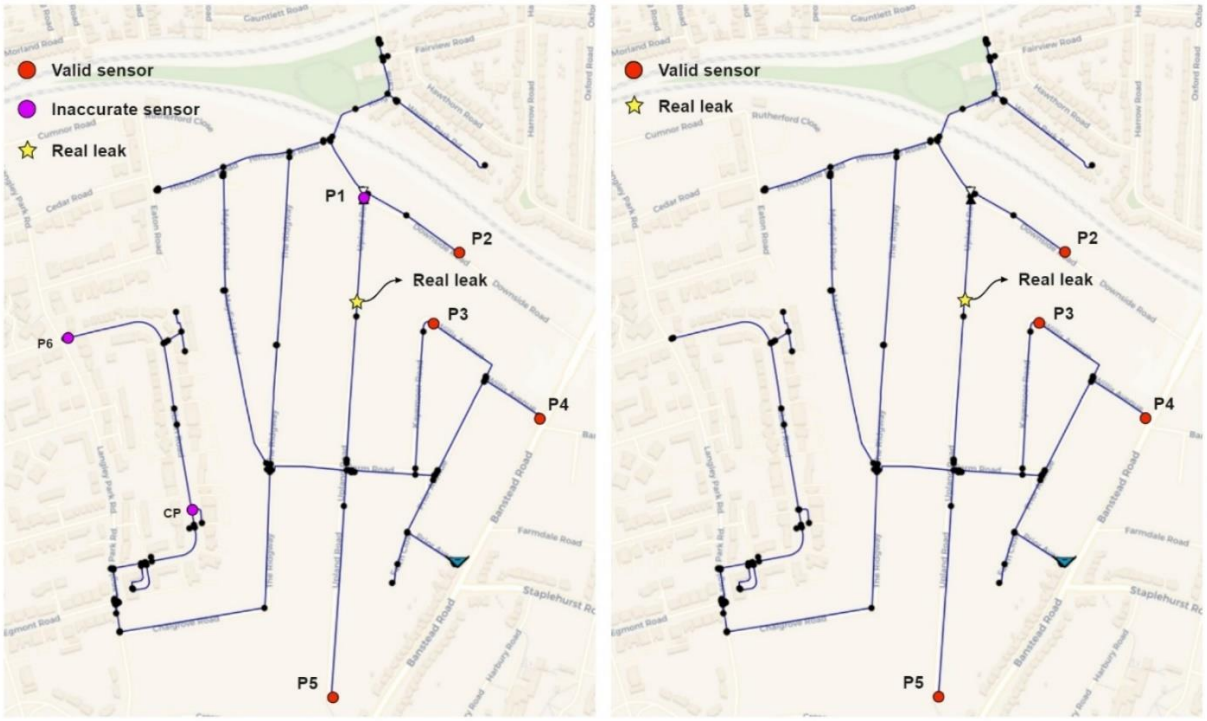


**Figure 21.** Distribution of sensors in DMA6. Left: all sensors are shown; Right: only four valid sensors are shown.

Leak events in DMA10 have the worst results as reported in Table 9. Although the results of FP path are acceptable, the value of LE is high, which indicates the long distance between the actual leak and the predicted leak node. DMA10 is the largest DMA with an average flow of 11.68 l/s, and it has a hybrid structure network with several loops on both sides of the transport main as shown in Figure 22. There are nine pressure sensors and P5 was removed as a faulty sensor. However, leak flow ranges from 7% to 15% of average flow. Pressure responses to leaks are very limited at about 0.2 m. These results are consistent with the fact that leak-driven pressure changes can be hidden by pressure variations due to customer use in larger DMAs. In addition, eight valid sensors concentrate in the northern area. If a leak is detected in the south, it will be hard to have a good prediction due to weak responses in pressure sensors.


**Figure 22.** The network layout of DMA10 and the location of three fire hydrant leaks.

In addition, another finding from Table 9 is that FP path is usually smaller than LE. Two metrics focus on candidate nodes and pipe length, respectively. The larger differences between LE and FP path values mainly appear in DMA2, DMA6, DMA8, and DMA10. A common feature of these leak locations in those four DMAs is that they are in areas where there is a relatively higher concentration of nodes in the DMA. Another reason is that some nodes inside the circle are probably not hydraulically related to others. Consequently, it seems that FP path has a higher tolerance for uneven distribution of nodes in the DMAs.

### 4.3.2 Results of model-based method in different situations

In Table 9, the results of leak events are divided into different DMAs for a summary. The performance is very different in different DMAs even with the same strategy. It is of interest to consider all leak events into different categories and to investigate the impact factors of the performance. According to results in section 4.3.1, model output correction and faulty sensor removal (MP+RE) work well in most cases. Therefore, this section focuses on the strategy of MP+RE.

1. The time of the leak

To investigate the relationship between the localization performance and the time of the leak, all regular leak events of 0.8 l/s or 1.5 l/s are categorized as the time of the day, morning, daytime, and nighttime. The average results in different time classes are reported in Table 10 below.

**Table 10.** The localization performance of MP+RE strategy in different time classes.

| Time of the leak | Leak events | MP+RE | |
|---|---|---|---|
| | | LE | FP path |
| Morning | 24 | 24.8% | 12.1% |
| Daytime | 10 | 25.2% | 12.7% |
| Nighttime | 4 | 28.5% | 21.0% |

It shows that the localization performance is better for leak events in the morning in terms of both LE and FP path, in comparison to the leaks that happened during the daytime or nighttime. The main reason is the more significant pressure responses to the leak (e.g., In DMA3, pressure drop driven by the leak in the morning: 0.71 m; daytime: 0.60 m; nighttime: 0.33 m). The number of leak events is unbalanced in different time classes, but the trend of performance is clear from morning to night. Although larger uncertainties exist in the morning time due to greater fluctuations in flow, the higher head loss still makes it easier to find the leak in the WDN (Moors et al., 2018).

2. Leak size

The leak flow of most leak events is set at 0.8 l/s or 1.5l/s. In addition, there are leak events at 2.0 l/s and 2.5 l/s in DMA3 and DMA4 (see section 3.2). Although leak flow is fixed and controlled by the fire hydrant, the relative leak size is different relative to each DMA. The relationship between the performance of MP+RE and relative leak size is explored in this section. Therefore, the results of all leak events are classified into three categories based on relative leak size: small, medium, and large. In Bakker et al. (2014), the authors propose a function of the average flow to estimate the minimum detectable leak size in the DMA:

$$Q_{leak,min} = 0.27 \times Q_{DMA}^{0.87} \quad (27)$$

which is about 20% of the average flow transformed into our case studies. As a result, the small class of relative leak size is set as 0% to 20%. It is useful to know to what extent the MP+RE strategy can reduce the search area of the hard-to-detect leak.

**Table 11.** The localization performance of MP+RE strategy in different classes of relative leak size.

| Relative leak size | Leak events | MP+RE | |
|---|---|---|---|
| | | LE | FP path |
| Small [0-20%] | 14 | 33.9% | 12.9% |
| Medium [20-40%] | 19 | 20.2% | 12.9% |
| Large [>40%] | 14 | 9.7% | 8.6% |

Table 11 shows the localization performance of the MP+RE strategy in different classes of relative leak flow. The average results in LE and FP path of 14 leak events fall into the 'good' class when the relative leak flow is larger than 40%. For the 19 medium-sized leak events, only two of them in DMA6 have poor results which have been explained in section 4.3.1. The rest of the leak events have 'fair' or 'good' results, which means the search area is reduced to less than 20% of the nodes or networks. For the 14 small-sized leak events, although the performance of localization is fair from the FP path point of view, there are three leak events in DMA10 where the results are very poor in terms of LE which are even higher than 50%. The relative leak flow of these three leaks is even smaller 10%, which are usually considered background leaks that are hard to detect and find.

These findings illustrate that the MP+RE strategy can effectively deal with large and most medium leaks. For small-size leaks with relative leak size > 10%, in general, a 'fair' or 'acceptable' result can be expected. Leaks with a relative leak size < 10% are difficult to locate by using model-based method with model output correction and fault sensor removal.

3. The location of the leak

In Pérez et al. (2011), the authors mention that the leak-driven pressure drops are not significant in a highly looped network as in their case study DMA Nova Icaria. In this study, DMA7 is a hybrid network selected to research the relationship between the leak location and localization performance. Six leak events were simulated at three different locations in DMA7 as shown in Figure 23.

**Table 12.** The localization results of MP+RE strategy in DMA7.

| DMA7 | Leak events | Leak location | Pressure drops [m] | MP+RE | |
|---|---|---|---|---|---|
| | | | | LE | FP path |
| Location 1 | 2 | Loop | 0.03 | 12.0% | 8.5% |
| Location 2 | 2 | Loop | 0.03 | 20.5% | 13.5% |
| Location 3 | 2 | Branch | 0.06 | 7.5% | 7.0% |

In Table 12, pressure drops mean the mean pressure drops at all sensor node caused by the leak. It is proved that pressure responses to the leak in branch area is more significant than that in loop area. As a result, the performance of the methodology in different areas is consistent with the magnitude of leak-driven pressure drops in different network structure areas.

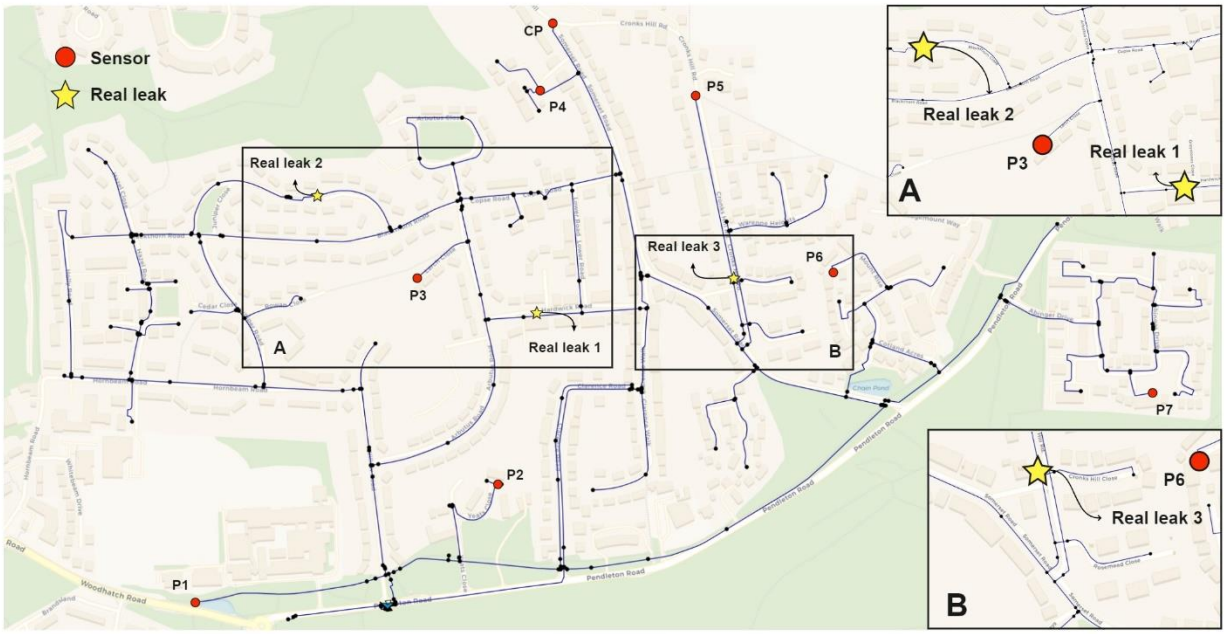


**Figure 23.** The network layout of DMA7. Three leak nodes are shown as yellow stars.

4. Scenario analysis time step

The scenario analysis time step $T_a$ is usually restricted by the sampling rate in practice. The initial setting of $T_a$ is 15 min in this study since the sampling rate of flow and pressure is usually 15 min (Bakker, Vreeburg, et al., 2014b; Romano et al., 2017b). In this research, both pressure and flow data are collected every 1 min. Therefore, the relationship between Ta and the localization results is investigated for 1 min, 5 min, 10 min, 15 min, and 30 min to see if it is worthwhile for water utilities to increase the sampling rate of sensors. In each DMA, two regular leak events of 0.8 l/s and 1.5 l/s are selected for this study.

**Table 13.** LE of MP+RE strategy for different scenario analysis time step

| LE | Leak events | Scenario analysis time step $T_a$ | | | | |
|---|---|---|---|---|---|---|
| DMA | | 30 min | 15 min | 10 min | 5 min | 1 min |
| DMA1 | 2 | 2.0% | 2.0% | 5.5% | 5.5% | 5.5% |
| DMA2 | 2 | 41.0% | 41.0% | 42.0% | 41.0% | 42.0% |
| DMA3 | 2 | 1.5% | 3.0% | 11.0% | 15.0% | 19.0% |
| DMA4 | 2 | 33.0% | 33.0% | 33.0% | 33.0% | 33.0% |
| DMA5 | 2 | 20.0% | 20.0% | 24.0% | 22.0% | 20.0% |
| DMA6 | 2 | 13.0% | 13.0% | 13.0% | 13.0% | 13.0% |
| DMA7 | 2 | 12.0% | 12.0% | 12.0% | 12.0% | 13.5% |
| DMA8 | 2 | 61.0% | 18.0% | 41.0% | 33.5% | 33.5% |
| DMA10 | 2 | 75.0% | 76.3% | 77.5% | 74.0% | 80.5% |
| Average | | 28.7% | 24.3% | 28.8% | 27.7% | 28.9% |

**Table 14.** FP path of MP+RE strategy for different scenario analysis time step

| FP path | Leak events | Scenario analysis time step $T_a$ | | | | |
|---|---|---|---|---|---|---|
| DMA | | 30 min | 15 min | 10 min | 5 min | 1 min |
| DMA1 | 2 | 6.0% | 4.5% | 5.0% | 5.0% | 5.0% |
| DMA2 | 2 | 40.5% | 32.0% | 76.5% | 50.0% | 55.5% |
| DMA3 | 2 | 8.0% | 1.0% | 13.5% | 15.5% | 15.5% |
| DMA4 | 2 | 37.0% | 39.0% | 37.0% | 37.0% | 37.0% |
| DMA5 | 2 | 10.0% | 8.0% | 12.0% | 8.0% | 8.0% |
| DMA6 | 2 | 2.0% | 2.0% | 2.0% | 2.0% | 2.0% |
| DMA7 | 2 | 9.0% | 8.5% | 9.0% | 8.0% | 9.0% |
| DMA8 | 2 | 10.0% | 1.0% | 8.0% | 20.5% | 22.0% |
| DMA10 | 2 | 18.0% | 25.0% | 21.0% | 20.0% | 20.0% |
| Average | | 15.6% | 13.4% | 20.4% | 18.4% | 19.3% |

Table 13 and Table 14 show the results for different scenario analysis time steps in terms of LE and FP path, respectively. The selection of Ta has an impact on the localization performance in most cases. There is not much difference in results between using 1-, 5- and 10-min solutions. Overall, their results are worse than those using 15- and 30-min solutions. This is most likely caused by the effect of the nuisance noise, in which case sensor measurements are not sufficiently smoothed out during the $T_a$ (Meseguer et al., 2014b). For the rest of solutions, the performance using 15-min solution is slightly better than that of a 30-min solution. It can be seen that a long $T_a$ is also not a good choice since subtle pressure changes caused by the leak may be filtered out, leading to worse results.

In Mounce et al. (2012), earlier detection of leaks can be achieved by a shorter sampling time. However, for the model-based leak localization method used in this research, a shorter sampling time can shorten the scenario analysis time step, but it provides a limited improvement in the localization performance, and a too-short $T_a$ can even result in worse results. In addition, a high sampling rate usually means extra costs for DMA instruments. Considering the results of different scenario analysis time steps, 15-min solution outperforms the others and is set as the $T_a$ for subsequent studies on interpolation and data-driven methods.

### 4.3.3 Model-based method with interpolation

In addition to graph-based interpolation, sensor calibration and inaccurate sensor removal are also included for comparison to investigate the effect of interpolation on the model-based leak localization. Based on the interpolation results of the synthetic data from EPANET simulation, the head loss weight and 50% of the number of sensors are set as the graph weight and frequency limit for graph-based interpolation, respectively.

The performance of model-based method with interpolation is reported in Table 15. It shows the percentage of leak events corresponding to each of the four result classes under each strategy. BASE stands for no model output correction and faulty sensors removal; RE stands for removing faulty sensors; MP means model output correction; GP means graph-based interpolation.

Table 15. The localization results of the percentage of leak events in different result classes.

| Results category | BASE+GP | | RE+GP | | MP+GP | | MP+RE+GP | |
|---|---|---|---|---|---|---|---|---|
| | LE | FP path | LE | FP path | LE | FP path | LE | FP path |
| Good [0-10%] | 21.3% | 31.9% | 4.3% | 12.8% | 19.1% | 12.8% | 8.5% | 10.6% |
| Fair [10-20%] | 10.6% | 4.3% | 6.4% | 8.5% | 10.6% | 8.5% | 14.9% | 10.6% |
| Acceptable [20-40%] | 29.8% | 34.0% | 25.5% | 25.5% | 25.5% | 29.8% | 25.5% | 12.8% |
| Poor [>40%] | 38.3% | 29.8% | 63.8% | 53.2% | 44.7% | 48.9% | 51.1% | 66.0% |

In general, with virtual sensors created by graph-based interpolation, the model-based method performs best in the absence of model output correction and faulty sensor removal (BASE+GP). 22.2% and 31.9% of leak events have good results in terms of LE and FP path, respectively. However, for the other strategies, more than half of leak events have poor results.

For four different strategies, RE+GP shows the worst performance, which is the same when there is no graph-based interpolation involved (see section 4.3.1). It can be explained by the lower accuracy of graph-based interpolation caused by fewer sensors (X. Zhou et al., 2022b). When interpolation is applied in practice using real data measurements, it differs from the behavior of WDNs in the simulated environment although the head loss weight obtained from model simulations is used to represent the hydraulic relationships of different nodes. The accuracy of interpolation is not quantifiable in terms of MAE or other statistical values due to the absence of data at nodes without sensors. As a result, it is difficult to compare the magnitude of errors of interpolation with pressure drops driven by the leak.

Comparing the results using interpolation with (BASE+GP) or without model output correction (MP+GP), correction even makes the performance even worse in general. The possible reason is the inappropriate use of model output correction tactics. For graph-based interpolation, the relationship of head signals between different nodes is interconnected by the head loss weight obtained from the EPANET simulator. However, with model output correction, head signals at each sensor node are changed by adding an offset value. Therefore, in this case, the hydraulic relationship between head signal at each node cannot be very precisely represented by the head loss weight.

Considering both LE and FP path, poor results (> 40%) are in DMA5, DMA7, DMA8 and DMA10 as reported in Table 16. Test flow in these DMAs is smaller with the relative leak flow less than 40% of the average flow. Apart from the smaller leak flow, another possible reason is the lower spatial density of sensors in the DMA. The ratio of total nodes to the number of pressure sensors is over 30 in these four DMAs. Compared with the results reported in Table 9, it can be seen that the negative effect of the low density of sensor on the localization performance is more significant when interpolation is used.

By comparing the average value of LE or FP path using different strategies, the overall performance of the model-based method with graph-based interpolation is worse than that without interpolation, which is especially true in the presence of model output correction as reported in Table 16.

Furthermore, graph-based interpolation requires more extra effort in computation. Taking the duration of leak tests of 0.8 l/s in this study as an example, the scenario analysis time step is 15 min, which means the interpolation needs to be done 5 times for a 1.5-hour leak for each node. The number of total nodes in 9 DMAs varies from 88 to 343, which is relatively small compared with the DMA Nova Icaria which is commonly used as a case study for leak localization. The computation work is even more greatly increased for large WDNs. As a result,

both in terms of improving the localization performance and the complexity of the approach itself, the interpolation cannot bring more benefits to the model-based leak localization. So far, based on the results of this study, interpolation is not recommended for integration with the model-based approach.

**Table 16.** Comparison of the performance of the best strategy with or without interpolation

| DMA | Leak events | Nodes/Sensors [-] | BASE | | BASE+GP | | MP+RE | |
|---|---|---|---|---|---|---|---|---|
| | | | LE | FP path | LE | FP path | LE | FP path |
| DMA1 | 4 | 15 | 55.8% | 23.5% | 34.5% | 15.5% | 1.0% | 2.3% |
| DMA2 | 2 | 23 | 9.0% | 96.0% | 37.5% | 26.0% | 41.0% | 32.0% |
| DMA3 | 10 | 16 | 37.6% | 26.8% | 17.7% | 17.1% | 5.8% | 5.6% |
| DMA4 | 7 | 16 | 53.7% | 42.4% | 27.9% | 28.7% | 17.9% | 14.9% |
| DMA5 | 4 | 31 | 32.0% | 7.5% | 34.3% | 47.3% | 30.0% | 24.0% |
| DMA6 | 6 | 18 | 18.5% | 2.1% | 35.0% | 19.7% | 40.0% | 13.7% |
| DMA7 | 6 | 36 | 35.3% | 49.0% | 43.0% | 45.7% | 13.33% | 7.83% |
| DMA8 | 2 | 43 | 60.5% | 30.0% | 47.5% | 60.0% | 18.0% | 1.0% |
| DMA10 | 6 | 37 | 75.0% | 56.3% | 73.8% | 55.5% | 43.0% | 12.7% |
| Average | | | 41.9% | 37.1% | 39.0% | 35.0% | 23.3% | 12.7% |

## 4.4 Data-driven leak localization

### 4.4.1 General results of data-driven method

To investigate the performance of the data-driven approach described in Section 2.3, seven DMAs with more than one regular leak test are selected (43 leak events). The three days of historical data in the non-leak situation are used to create the pressure interpolation maps to be compared with the maps created during the leak. As mentioned earlier, the data are reduced to every 15 min by averaging all the samples. As a result, the method is applied to each analysis time step (15 min) to obtain a residual map. Finally, taking the evolution of the residual maps into account, Bayesian temporal reasoning is used to generate an overall result during the leak event and give a prediction on the leak node with the highest possibility.

Although the accuracy of interpolation is the best using the model-simulated head loss as the graph weight (in section 4.2), the uniform weight is applied for the data-driven method in order to minimize the dependence on the network information. Here only the connectivity and typology of the network are required for graph-based interpolation. As described in section 2.1.3 and 2.1.4, both model output correction and faulty sensor detection require more network information. Therefore, the original data is used for data-driven leak localization. The residual map is generated at each time step and the overall result of each leak event will be obtained using Bayesian temporal reasoning, identifying the most probable leak node. Due to the difference in principles of leak localization, the performance of the data-driven approach is evaluated by localization error (LE).

The performance of the data-driven approach is shown in Table 17. It shows the percentage of leak events corresponding to each of the four result classes under each strategy. LE stands for localization error. Overall, 30.2% of leak events have good results with the candidate nodes less than 10% of nodes in the DMA. Less than 18.6% of leak events have poor results (> 40%). The average value of LE of all leak events is 23.8% as reported in Table 18, which is even similar to the average results (LE=23.6%) of the model-based method with model output correction and faulty sensor removal.

Table 17. The localization results of the percentage of leak events in different result classes.

| Results category | Data-driven |
|---|---|
| | LE |
| Good [0-10%] | 30.2% |
| Fair [10-20%] | 27.9% |
| Acceptable [20-40%] | 23.3% |
| Poor [>40%] | 18.6% |

The localization results of leak events using the data-driven method are shown in Table 18. Localization error (LE) ranges from 6.0% to 48.3% in different DMAs. It is very clear that LE is strongly related to the spatial density of pressure sensors quantified by the ratio of the number of nodes to sensors. To be more specific, the correlation between LE and the ratio of nodes to sensors is extremely high as 0.957. This is most likely explained by the higher accuracy of graph-based interpolation in the presence of more sensors. It was supported by the research from Zhou et al. (2022), in which the authors used synthetic data in a real middle-sized DMA. In this way, it can be seen that the effect of the density of sensors on leak localization performance is consistent with its effect on interpolation precision when graph-based interpolation is applied. In addition, a relatively high correlation of 0.732 also exists between LE and the average flow of the DMA. It can be attributed to the small proportion of leak flow to the average flow. This phenomenon is similar to the localization results using the model-based method.

Table 18. Performance of data-driven method in different DMAs.

| DMA | Leak events | Average flow [l/s] | Nodes/ Sensors [-] | Data-driven |
|---|---|---|---|---|
| | | | | LE |
| DMA1 | 4 | 1.4 | 15 | 12.0% |
| DMA3 | 10 | 2.4 | 16 | 13.2% |
| DMA4 | 7 | 3.8 | 16 | 6.0% |
| DMA5 | 4 | 3.8 | 31 | 26.5% |
| DMA6 | 6 | 4.3 | 18 | 16.0% |
| DMA7 | 6 | 5.5 | 36 | 48.3% |
| DMA10 | 6 | 11.7 | 37 | 44.7% |
| Average | | | | 23.8% |

In Romero-Ben et al. (2022), the authors used graph-based interpolation with the length weight (see section 2.2.1) for pressure head estimation and a different approach for localization analysis and tested it using the synthetic data from the BattLeDIM 2020 challenge. The benchmark network comprises three areas, where Area A contains 655 nodes with 29 pressure sensors (the ratio of nodes to sensors is 22). The implementation of their data-driven method helps to locate 8 leak events in Area A. 6 out of 8 synthetic leaks are located within 300 m away from the actual leak pipes. Taking the DMA size and difference in datasets into account, this is comparable to our findings in the DMAs with a similar level of the density of sensors as shown in Table 19. This is the first work for the implementation of data-driven method with graph-based interpolation for leak localization using real data measurements.

**Table 19.** Comparison of localization performance in Romero-Ben et al. (2022) and this study.

| Area A in L-Town | Distance [m] | DMA in the UK | Distance [m] |
|---|---|---|---|
| Leak 1 | 299.33 | Leak 1 | 339 |
| Leak 2 | 192.88 | Leak 2 | 67 |
| Leak 3 | 335.93 | Leak 3 | 377 |
| Leak 4 | 249.28 | Leak 4 | 64 |
| Leak 5 | 278.91 | Leak 5 | 81 |
| Leak 6 | 435.27 | Leak 6 | 171 |
| Leak 7 | 224.31 | | |
| Leak 8 | 74.28 | | |

As discussed in section 4.2, head loss weight achieves the highest accuracy of graph-based interpolation with the synthetic data without noise. However, in this section, the uniform weight is set initially as the graph weight for interpolation since it needs the least information from the model network. Length weight, Darcy weight, and head loss weight progressively increase the requirements of the model parameters (see Table 20). To further explore the relationship between localization performance and different graph weights, one leak event in each of 7 DMAs is selected.

**Table 20.** The model information is required for each graph weight calculation.

| Graph weight | Topology | Pipe length | Pipe diameter | Pipe roughness |
|---|---|---|---|---|
| Uniform weight | + | - | - | - |
| Length weight | + | + | - | - |
| Darcy weight | + | + | + | - |
| Head loss weight | + | + | + | + |

*The symbol '+' means the information is required.

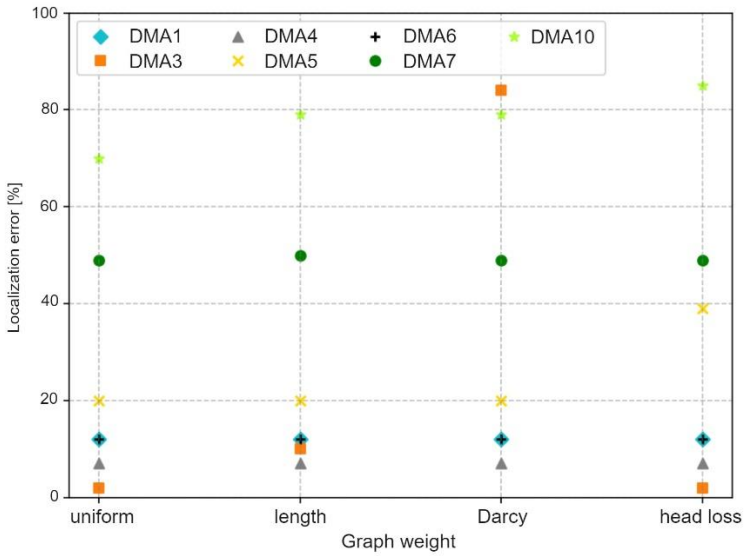


**Figure 24.** The relationship between the graph weight and localization error.

From Figure 24, the uniform weight slightly outperforms the others. There is not much difference between the results of uniform weight and head loss weight, except for the leak events in DMA5 and DMA10. Generally, it seems that head loss weight with more model information does not bring better localization results. The possible reason is the inaccurate model parameters in DMA5 and DMA10 and there is no strong proof of the interpolation accuracy in real life. Therefore, uniform weight is still recommended for this data-driven localization approach.

### 4.4.2 Comparison to model-based method

Compared with model-based leak localization, the data-driven method is less dependent on accurate information about WDNs, which is the biggest advantage of this method. Apart from the accurate model parameters, the model-based approach requires leak flow, which is difficult to know exactly in real life unless an algorithm for leak flow estimation has been implemented. As a result, in cases where the two approaches perform similarly or where no sufficient model parameters are available, the data-driven approach is theoretically better.

43 leak events in 7 DMAs were analyzed by both model-based and data-driven approaches. Therefore, the results of these leak tests are compared in this section. As shown in section 4.4.1, apart from the small leak flow, the density of sensors is an important impact factor. Localization error (LE) is used to evaluate the localization performance. Table 21 shows the performance of the data-driven method and model-based method. BASE stands for the model-base method without correction or faulty sensors removal. MP+RE means the model-based method with model output correction and faulty sensor removal.

**Table 21.** General comparison of data-driven method and model-based method.

| DMA | Leak events | Average flow [l/s] | Nodes/Sensors [-] | Data-driven | MP+RE | BASE |
|---|---|---|---|---|---|---|
| | | | | LE | LE | LE |
| DMA1 | 4 | 1.4 | 15 | 12.0% | 1.0% | 55.8% |
| DMA3 | 10 | 2.4 | 16 | 13.2% | 5.8% | 37.6% |
| DMA4 | 7 | 3.8 | 16 | 6.0% | 17.9% | 53.7% |
| DMA5 | 4 | 3.8 | 31 | 26.5% | 30.0% | 32.0% |
| DMA6 | 6 | 4.3 | 18 | 16.0% | 40.0% | 18.5% |
| DMA7 | 6 | 5.5 | 36 | 48.3% | 13.3% | 35.3% |
| DMA10 | 6 | 11.7 | 37 | 44.7% | 43.0% | 75.0% |
| Average | | | | 23.8% | 21.6% | 44.0% |

Without model output correction and faulty sensor removal, the data-driven method outperforms the model-based method in every DMAs. With model output correction and faulty sensor removal, the average results of the model-based method improve a lot, which is better in DAM1, DMA3, DMA7 and DMA10. MP+RE strategy is more robust and general in different situations, especially in the case of limited sensors. However, the comparison of the results using three strategies in Table 21 also illustrates the high dependence of the model-based method on the accuracy of the hydraulic model. In contrast, data-driven method can give ‘good’ or ‘fair’ results with a rough model with only typology information in the presence of more sensors.

Table 22. Results of long-time leak events (6h).

| DMA | Test time [h] | Data-driven | MP+RE |
|---|---|---|---|
| | | LE | LE |
| DMA4 | 09:00-15:00 | 0.26 | 0 |

Another finding in long-time leak events is reported in Table 22. It shows the performance of long-time leak tests using both model-based and data-driven methods. The result shows that the data-driven approach performs better on long-time leak events in DMA4, where the maximum correlation node is the same as the actual leak. The possible reason could be the Bayesian temporal reasoning (see section 2.3) recursively during the 6 hours.

In summary, in DMA with high-density pressure sensors, the data-driven approach using graph-based interpolation has comparable or even better localization performance than the model-based approach with model output correction and faulty sensor removal. Combined the results in Romero-Ben et al. (2022) with the results in this research, the data-driven approach can be used in practice to reduce the search area of the leak in any case where the DMA is equipped with high-density sensors and the model parameters are insufficient or inaccurate. On the contrary, with fewer sensors and a high-quality hydraulic model, the model-based approach has better localization performance in more general conditions.

# Chapter 5: Conclusions and Recommendations

As the main contribution of non-revenue water, leakage is an important issue not only for drinking water companies but also for all people and society. This study presents the implementation of model-based method and data-driven method to solve the problem of leak localization for case studies in South London, UK. Both two methods are tested on artificial fire hydrant leak tests. On the one hand, the model-based method is used and the performance is highly improved with model output correction and faulty sensor removal. On the other hand, a data-driven method with graph-based interpolation is developed and proved to perform well in areas with a high spatial density of pressure sensors. The two main research questions proposed in section 1.4 are answered.

1. How can we improve the model-based leak localization relative to the method developed by Quevedo et al. (2011)?

The main idea of the model-based method proposed by Quevedo et al. (2011) is to calculate the correlation between the pressure drops driven by each potential leak in the model and the pressure drops in real WDNs. Fewer sensors and large model errors are two factors that have a negative impact on localization precision. To solve these two problems, model output correction and faulty sensor removal, as well as graph-based interpolation are introduced in this work.

Results from testing the method in fire hydrant leaks show that the strategy of model output correction and faulty sensor removal greatly improves the performance of the model-based method in the absence of highly accurate models. In this case, the pressure responses to the leak can be easier to recognize. The search area of 68.1% of leak events can be reduced by more than 80% with respect to total nodes in the DMA. Considering the pipes, 63.8% of leak events are possible to find within 10% of the total networks. In general, this strategy works reasonably well on detectable leaks (relative leak size > 20%) and also has the chance to locate hard-to-find leaks with acceptable results. It is noteworthy that the removal of faulty sensors should be done with great care, especially in areas with very limited sensors. In addition, other unfavorable conditions, including nighttime and loop structure correspond to lower pressure drops caused by the leak. Sensitivity analysis to scenario analysis time steps Ta shows that an excessively short Ta cannot enhance the localization precision due to the negative effect of measurement noise. However, if Ta is too long, the subtle pressure responses to the leak are probably filtered out, leading to poor results. The optimal analysis time step is 15 min in this work.

Graph-based interpolation cannot further improve the model-based leak localization. With interpolation, more than 50% of leak tests have poor results (> 40%) in terms of either LE or FP path. The localization performance is even worse when the interpolation and model output correction are implemented at the same time due to the change in the hydraulic relationship between different nodes.

For the wide application of the proposed methodology in practice, some research questions need to be investigated. For example, it is of interest to assess whether the quadratic function proposed in this work for model output correction can also work well in WDNs from other countries or using different types of pressure sensors and if there is another function that can better explain the relationship between flow and residuals. In addition, the criteria for faulty sensor detection are relatively specific in our cases. Consequently, more research is needed to complete the process of faulty pressure sensor detection to ensure that only the measurements from normal sensors are used for localization analysis.

Furthermore, the principle of either the LE or FP path proposed in this work or the performance metrics developed by other researchers is related to the information of the real leak location. However, in real life, the location of a leak is unknown to us. Therefore, it is very useful to find an indicator strongly correlated to the reliability of leak localization. In this case, water companies can make a better decision on whether to trust the localization results and take action.

2. Can we advise a data-driven approach using pressure and graph-based interpolation for leak localization in real WDNs?

In this work, the data-driven method is proposed to locate leaks. Graph-based interpolation is used to create pressure maps in non-leak and leak situations. Compared with the model-based method used in this study, the implementation of the data-driven method for leak localization gives two main advantages. Firstly, it is less dependent on the model accuracy since it only requires topology information of the network. Secondly, the volume of leak flow is not required for this methodology. As we know, it is hard to measure in real life unless an effective method is applied for the calculation/estimation. The use of graph-based interpolation for head signal estimation and for data-driven leak localization has been investigated (see section 2.3) in real WDNs and benchmark networks, respectively. However, this is the first time that the data-driven method involving graph-based interpolation is tested on fire hydrant leak tests using real data measurements.

This methodology performs fairly well in fire hydrant leaks. For 58.1% of leak events, the search area involved less than 20% of all nodes. Results of leak events in different DMAs show how important the spatial density of pressure sensors is to localization precision. Combing the results in Romero-Ben et al. (2022) and in this study, this methodology is expected to give sufficiently accurate results in areas with a node-to-sensor ratio of about 20.

Compared with the results of the model-based method with model output correction and faulty sensor removal, the data-driven method can give comparable or even better localization results in areas with a higher spatial density of pressure sensors. As a consequence, the data-driven approach is highly recommended in cases where the water utilities cannot provide accurate hydraulic model information or where the network parameters are not easily measured. However, when the data measurements are very limited, the model-based method shows better performance in general. Water companies need to weigh the cost of extra pressure sensors against the cost of investigating and measuring model parameters to determine which approach to eventually use in practice.

To further improve the performance of this data-driven method, several research tasks remain open. In our case, the deployment of pressure sensors (amount, placement) is more friendly to the model-based method. Therefore, it is interesting to investigate the optimal deployment of sensors aiming at improving the accuracy of graph-based interpolation. Since interpolation provides more available data in non-leak and leak situations, it is worthwhile to study how to process these data to identify the leak. The localization strategy used in this work proposed by Soldevila et al. (2019) is different from that in Romero-Ben et al. (2022). More research on localization analysis needs to be done after the interpolation stage. In addition, in this study, three days of historical data are used to generate the pressure map in the non-leak situation. A related task is to investigate the performance of this methodology using historical data of different days.

# Reference


Alves, D., Blesa, J., Cembrano, G., Puig, V., & Duviella, E. (2022). Leak Localization in Water Distribution Networks Using Data-Driven and Model-Based Approaches [Review of *Leak Localization in Water Distribution Networks Using Data-Driven and Model-Based Approaches*, by L. Romero-Ben]. *Journal of Water Resources Planning and Management*, *148*(5), 04022016. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001542

Bakker, M., Trietsch, E. A., Vreeburg, J. H. G., & Rietveld, L. C. (2014). Analysis of historic bursts and burst detection in water supply areas of different size. *Water Supply*, *14*(6), 1035–1044. https://doi.org/10.2166/ws.2014.063

Bakker, M., Vreeburg, J. H. G., Van De Roer, M., & Rietveld, L. C. (2014a). Heuristic burst detection method using flow and pressure measurements. *Journal of Hydroinformatics*, *16*(5), 1194–1209. https://doi.org/10.2166/hydro.2014.120

Bakker, M., Vreeburg, J. H. G., Van De Roer, M., & Rietveld, L. C. (2014b). Heuristic burst detection method using flow and pressure measurements. *Journal of Hydroinformatics*, *16*(5), 1194–1209. https://doi.org/10.2166/hydro.2014.120

Bentley Systems, Walski, T. M., Chase, D. V., Savic, D. A., Grayman, W., Beckwith, S., & Koelle, E. (2003). *Advanced Water Distribution Modeling and Management*. https://app.knovel.com/hotlink/toc/id:kpAWDMM009/advanced-water-distribution/advanced-water-distribution

Blesa, J., & Pérez, R. (2018). Modelling uncertainty for leak localization in Water Networks. *IFAC-PapersOnLine*, *51*(24), 730–735. https://doi.org/10.1016/j.ifacol.2018.09.656

Casillas Ponce, M. V., Garza Castañón, L. E., & Cayuela, V. P. (2014). Model-based leak detection and location in water distribution networks considering an extended-horizon analysis of pressure sensitivities. *Journal of Hydroinformatics*, *16*(3), 649–670. https://doi.org/10.2166/hydro.2013.019

Colombo, A. F., Lee, P., & Karney, B. W. (2009). A selective literature review of transient-based leak detection methods. *Journal of Hydro-Environment Research*, *2*(4), 212–227. https://doi.org/10.1016/j.jher.2009.02.003

El-Zahab, S., & Zayed, T. (2019). Leak detection in water distribution networks: An introductory overview. *Smart Water*, *4*(1), 5. https://doi.org/10.1186/s40713-019-0017-x

Farley, B., Mounce, S. R., & Boxall, J. B. (2013). Development and Field Validation of a Burst Localization Methodology. *Journal of Water Resources Planning and Management*, *139*(6), 604–613. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000290

Gao, Y., Brennan, M. J., Joseph, P. F., Muggleton, J. M., & Hunaidi, O. (2005). On the selection of acoustic/vibration sensors for leak detection in plastic water pipes. *Journal of Sound and Vibration*, *283*(3–5), 927–941. https://doi.org/10.1016/j.jsv.2004.05.004

Javadiha, M., Blesa, J., Soldevila, A., & Puig, V. (2019). Leak Localization in Water Distribution Networks using Deep Learning. *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, 1426–1431. https://doi.org/10.1109/CoDIT.2019.8820627

Klise, K. A., Hart, D. B., Moriarty, D., Bynum, M. L., Murray, R., Burkhardt, J., & Haxton, T. (2017). *Water Network Tool for Resilience (WNTR) User Manual*. 50.

Li, J., Zheng, W., & Lu, C. (2022). An Accurate Leakage Localization Method for Water Supply Network Based on Deep Learning Network. *Water Resources Management*, *36*(7), 2309–2325. https://doi.org/10.1007/s11269-022-03144-x

Li, R., Huang, H., Xin, K., & Tao, T. (2015). A review of methods for burst/leakage detection and location in water distribution systems. *Water Supply*, *15*(3), 429–441. https://doi.org/10.2166/ws.2014.131

Liemberger, R., & Wyatt, A. (2019a). Quantifying the global non-revenue water problem. *Water Supply*, *19*(3), 831–837. https://doi.org/10.2166/ws.2018.129

Liemberger, R., & Wyatt, A. (2019b). Quantifying the global non-revenue water problem. *Water Supply*, *19*(3), 831–837. https://doi.org/10.2166/ws.2018.129

Meseguer, J., Mirats-Tur, J. M., Cembrano, G., Puig, V., Quevedo, J., Pérez, R., Sanz, G., & Ibarra, D. (2014a). A decision support system for on-line leakage localization. *Environmental Modelling & Software*, *60*, 331–345. https://doi.org/10.1016/j.envsoft.2014.06.025

Meseguer, J., Mirats-Tur, J. M., Cembrano, G., Puig, V., Quevedo, J., Pérez, R., Sanz, G., & Ibarra, D. (2014b). A decision support system for on-line leakage localization. *Environmental Modelling & Software*, *60*, 331–345. https://doi.org/10.1016/j.envsoft.2014.06.025

Moors, J., Scholten, L., van der Hoek, J. P., & den Besten, J. (2018). Automated leak localization performance without detailed demand distribution data. *Urban Water Journal*, *15*(2), 116–123. https://doi.org/10.1080/1573062X.2017.1414272

Pérez, R., Puig, V., Pascual, J., Quevedo, J., Landeros, E., & Peralta, A. (2011). Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks. *Control Engineering Practice*, *19*(10), 1157–1167. https://doi.org/10.1016/j.conengprac.2011.06.004

Perez, R., Sanz, G., Puig, V., & Quevedo, J. (2014). Leak Localization in Water Networks: A Model-Based Methodology Using Pressure Sensors Applied to a Real Network in Barcelona [Applications of Control]. *IEEE Control Systems*, *34*(4), 24–36. https://doi.org/10.1109/MCS.2014.2320336

Preis, A., Whittle, A. J., Ostfeld, A., & Perelman, L. (2011). Efficient Hydraulic State Estimation Technique Using Reduced Models of Urban Water Networks. *Journal of Water Resources Planning and Management*, *137*(4), 343–351. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000113

Puust, R., Kapelan, Z., Savic, D. A., & Koppel, T. (2010). A review of methods for leakage management in pipe networks. *Urban Water Journal*, *7*(1), 25–45. https://doi.org/10.1080/15730621003610878

Quevedo, J., Cugueró, M. À., Pérez, R., Nejjari, F., Puig, V., & Mirats, J. M. (2011). *Leakage Location in Water Distribution Networks based on Correlation Measurement of Pressure Sensors*. 9.

Quiñones-Grueiro, M., Bernal-de Lázaro, J. M., Verde, C., Prieto-Moreno, A., & Llanes-Santiago, O. (2018). Comparison of Classifiers for Leak Location in Water Distribution Networks. *IFAC-PapersOnLine*, *51*(24), 407–413. https://doi.org/10.1016/j.ifacol.2018.09.609

Romano, M., Kapelan, Z., & Savić, D. A. (2013). Geostatistical techniques for approximate location of pipe burst events in water distribution systems. *Journal of Hydroinformatics*, *15*(3), 634–651. https://doi.org/10.2166/hydro.2013.094

Romano, M., Kapelan, Z., & Savić, D. A. (2014). Evolutionary Algorithm and Expectation Maximization Strategies for Improved Detection of Pipe Bursts and Other Events in Water Distribution Systems. *Journal of Water Resources Planning and Management*, *140*(5), 572–584. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000347

Romano, M., Woodward, K., & Kapelan, Z. (2017a). Statistical Process Control Based System for Approximate Location of Pipe Bursts and Leaks in Water Distribution Systems. *Procedia Engineering*, *186*, 236–243. https://doi.org/10.1016/j.proeng.2017.03.235

Romano, M., Woodward, K., & Kapelan, Z. (2017b). Statistical Process Control Based System for Approximate Location of Pipe Bursts and Leaks in Water Distribution Systems. *Procedia Engineering*, *186*, 236–243. https://doi.org/10.1016/j.proeng.2017.03.235

Romero-Ben, L., Alves, D., Blesa, J., Cembrano, G., Puig, V., & Duviella, E. (2022). Leak Localization in Water Distribution Networks Using Data-Driven and Model-Based Approaches. *Journal of Water Resources Planning and Management*, *148*(5), 04022016. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001542

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

Soldevila, A., Blesa, J., Fernandez-Canti, R. M., Tornil-Sin, S., & Puig, V. (2019a). Data-Driven Approach for Leak Localization in Water Distribution Networks Using Pressure Sensors and Spatial Interpolation. *Water*, *11*(7), 1500. https://doi.org/10.3390/w11071500

Soldevila, A., Blesa, J., Fernandez-Canti, R. M., Tornil-Sin, S., & Puig, V. (2019b). Data-Driven Approach for Leak Localization in Water Distribution Networks Using Pressure Sensors and Spatial Interpolation. *Water*, *11*(7), 1500. https://doi.org/10.3390/w11071500

Soldevila, A., Blesa, J., Jensen, T. N., Tornil-Sin, S., Fernandez-Canti, R. M., & Puig, V. (2021). Leak Localization Method for Water-Distribution Networks Using a Data-Driven Model and Dempster–Shafer Reasoning. *IEEE Transactions on Control Systems Technology*, *29*(3), 937–948. https://doi.org/10.1109/TCST.2020.2982349

Soldevila, A., Blesa, J., Tornil-Sin, S., Duviella, E., Fernandez-Canti, R. M., & Puig, V. (2016). Leak localization in water distribution networks using a mixed model-based/data-driven approach. *Control Engineering Practice*, *55*, 162–173. https://doi.org/10.1016/j.conengprac.2016.07.006

Soldevila, A., Fernandez-Canti, R. M., Blesa, J., Tornil-Sin, S., & Puig, V. (2017). Leak localization in water distribution networks using Bayesian classifiers. *Journal of Process Control*, *55*, 1–9. https://doi.org/10.1016/j.jprocont.2017.03.015

Sophocleous, S., Savić, D., & Kapelan, Z. (2019). Leak Localization in a Real Water Distribution Network Based on Search-Space Reduction. *Journal of Water Resources Planning and Management*, *145*(7), 04019024. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001079

Stanković, L., Daković, M., & Sejdić, E. (2019). Introduction to Graph Signal Processing. In *Vertex-Frequency Analysis of Graph Signals* (pp. 3–108). Cham : Springer International Publishing : Springer; WorldCat.org. https://doi.org/10.1007/978-3-030-03574-7_1

University of Sheffield. (2015). *Leaky pipes can allow contaminants into our drinking water*. https://www.sheffield.ac.uk/news/nr/leaky-water-pipes-contaminants-drinking-water-1.470248

WHO. (2019). *WHO report*. https://www.who.int/news/item/18-06-2019-1-in-3-people-globally-do-not-have-access-to-safe-drinking-water-unicef-who

Wu, Y., & Liu, S. (2017). A review of data-driven approaches for burst detection in water distribution systems. *Urban Water Journal*, *14*(9), 972–983. https://doi.org/10.1080/1573062X.2017.1279191

Wu, Y., Liu, S., Wu, X., Liu, Y., & Guan, Y. (2016). Burst detection in district metering areas using a data driven clustering algorithm. *Water Research*, *100*, 28–37. https://doi.org/10.1016/j.watres.2016.05.016

Zaman, D., Tiwari, M. K., Gupta, A. K., & Sen, D. (2020). A review of leakage detection strategies for pressurised pipeline in steady-state. *Engineering Failure Analysis*, *109*, 104264. https://doi.org/10.1016/j.engfailanal.2019.104264

Zhang, Q., Wu, Z. Y., Zhao, M., Qi, J., Huang, Y., & Zhao, H. (2016). Leakage Zone Identification in Large-Scale Water Distribution Systems Using Multiclass Support Vector Machines. *Journal of Water Resources Planning and Management*, *142*(11), 04016042. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000661

Zhou, S. L., McMahon, T. A., Walton, A., & Lewis, J. (2002). Forecasting operational demand for an urban water supply zone. *Journal of Hydrology*, *259*(1–4), 189–202. https://doi.org/10.1016/S0022-1694(01)00582-0

Zhou, X., Liu, S., Xu, W., Xin, K., Wu, Y., & Meng, F. (2022a). Bridging hydraulics and graph signal processing: A new perspective to estimate water distribution network pressures. *Water Research*, *217*, 118416. https://doi.org/10.1016/j.watres.2022.118416

Zhou, X., Liu, S., Xu, W., Xin, K., Wu, Y., & Meng, F. (2022b). Bridging hydraulics and graph signal processing: A new perspective to estimate water distribution network pressures. *Water Research*, *217*, 118416. https://doi.org/10.1016/j.watres.2022.118416

# Appendix 1. Sensor evaluation

Table 23 shows the properties of all pressure sensors in our case studies. Orange color indicates the faulty sensors.

**Table 23.** Sensor evaluation

| SensorID | MAPE_0 [%] | C1 | C2 | MAPE_1 [%] |
|---|---|---|---|---|
| DMA2_P1 | 0.02 | 15767.36 | -0.08 | 0.01 |
| DMA2_P2 | 0.94 | 15068.10 | -0.94 | 0.01 |
| DMA2_P3 | 1.40 | 11147.25 | -1.37 | 0.01 |
| DMA2_P4 | 0.71 | -100142.56 | -0.32 | 0.15 |
| DMA2_P5 | 1.28 | -151875.63 | -0.66 | 0.17 |
| DMA2_CP | 0.56 | -124289.18 | 0.99 | 0.14 |
| | | | | |
| DMA5_P1 | 0.95 | 27786.62 | 0.41 | 0.07 |
| DMA5_P2 | 1.17 | 28031.66 | 0.62 | 0.07 |
| DMA5_P3 | 2.43 | 28212.24 | 1.79 | 0.07 |
| DMA5_P4 | 1.40 | 27732.33 | 0.84 | 0.07 |
| DMA5_P5 | 1.28 | 28199.85 | 0.71 | 0.07 |
| DMA5_P6 | 1.16 | 27846.03 | 0.60 | 0.07 |
| DMA5_CP | 2.51 | 27317.01 | 1.93 | 0.07 |
| | | | | |
| DMA8_P1 | 0.43 | 2464.17 | 0.33 | 0.10 |
| DMA8_P2 | 0.23 | 8803.43 | -0.76 | 0.12 |
| DMA8_P3 | 0.42 | 1265.08 | 0.39 | 0.11 |
| DMA8_P4 | 0.15 | 1248.98 | 0.09 | 0.11 |
| DMA8_P5 | 0.11 | 798.48 | 0.07 | 0.14 |
| DMA8_P6 | 0.08 | 338.57 | -0.11 | 0.11 |
| DMA8_P7 | 0.34 | 285.99 | 0.36 | 0.12 |
| DMA8_CP | 0.67 | 13861.30 | -1.53 | 0.11 |
| | | | | |
| DMA1_P1 | 0.32 | 40415.99 | -0.53 | 0.01 |
| DMA1_P2 | 0.26 | 19210.37 | -0.40 | 0.02 |
| DMA1_P3 | 0.22 | 67014.89 | 0.20 | 0.01 |
| DMA1_P4 | 0.11 | 68763.86 | -0.28 | 0.01 |
| DMA1_P5 | 0.26 | 65825.46 | -0.49 | 0.01 |
| DMA1_CP | 0.31 | 71866.67 | -0.57 | 0.01 |
| | | | | |
| DMA6_P1 | 2.22 | -30249.18 | -2.02 | 0.13 |
| DMA6_P2 | 1.89 | -16135.33 | -1.93 | 0.03 |
| DMA6_P3 | 1.72 | -10623.98 | -1.85 | 0.01 |
| DMA6_P4 | 1.89 | -10759.13 | -2.05 | 0.01 |
| DMA6_P5 | 1.73 | -10987.47 | -1.85 | 0.01 |
| DMA6_P6 | 2.73 | 2664.45 | -3.28 | 0.32 |
| DMA6_CP | 1.77 | -2155.99 | -2.07 | 0.25 |
| | | | | |
| DMA3_P1 | 0.94 | -118340.23 | -0.01 | 0.36 |

| | | | | |
|---|---|---|---|---|
| DMA3_P2 | 1.27 | -103481.26 | -0.39 | 0.12 |
| DMA3_P3 | 0.67 | -100059.42 | 0.09 | 0.09 |
| DMA3_P4 | 0.75 | -101121.85 | 0.03 | 0.09 |
| DMA3_P5 | 0.37 | -101464.62 | 0.35 | 0.09 |
| DMA3_CP | 1.84 | -97233.36 | 2.21 | 0.10 |
| | | | | |
| DMA10_P1 | 1.35 | 1446.92 | 0.63 | 0.07 |
| DMA10_P2 | 0.19 | 1641.37 | -0.14 | 0.04 |
| DMA10_P3 | 1.22 | 1787.09 | 0.50 | 0.04 |
| DMA10_P4 | 1.16 | 1766.22 | 0.46 | 0.04 |
| DMA10_P5 | 3.37 | -1176.72 | -1.92 | 0.78 |
| DMA10_P6 | 0.35 | 992.69 | 0.07 | 0.07 |
| DMA10_P7 | 0.72 | 1000.84 | -0.60 | 0.07 |
| DMA10_P8 | 1.72 | 2230.48 | 0.75 | 0.04 |
| DMA10_CP | 1.06 | 1224.70 | 0.49 | 0.03 |
| | | | | |
| DMA4_P1 | 22.69 | -284371.14 | 19.31 | 1.61 |
| DMA4_P2 | 1.53 | -14479.65 | -0.54 | 0.12 |
| DMA4_P3 | 1.74 | -14987.06 | -0.64 | 0.13 |
| DMA4_P4 | 3.10 | -11237.50 | -1.36 | 0.11 |
| DMA4_P5 | 0.11 | -8584.35 | 0.08 | 0.10 |
| DMA4_P6 | 0.54 | -11655.26 | -0.10 | 0.13 |
| DMA4_CP | 0.55 | -10744.52 | 0.45 | 0.14 |
| | | | | |
| DMA7_P1 | 0.98 | -1837.09 | -1.22 | 0.01 |
| DMA7_P2 | 0.74 | -2298.19 | -0.88 | 0.01 |
| DMA7_P3 | 0.99 | -19862.46 | -0.54 | 0.10 |
| DMA7_P4 | 1.14 | -3092.62 | -1.38 | 0.01 |
| DMA7_P5 | 0.01 | -2727.88 | 0.09 | 0.01 |
| DMA7_P6 | 3.15 | -2843.81 | -3.92 | 0.01 |
| DMA7_P7 | 0.34 | -3416.03 | -0.32 | 0.01 |
| DMA7_CP | 0.08 | -2769.74 | 0.01 | 0.00 |
| | | | | |
| DMA9_P1 | 0.48 | -9009.18 | 1.11 | 1.03 |
| DMA9_P2 | 0.46 | -8868.28 | 1.08 | 1.02 |
| DMA9_P3 | 0.76 | -7636.10 | 1.23 | 1.00 |
| DMA9_P4 | 0.15 | -6568.06 | 0.64 | 1.00 |
| DMA9_P5 | 1.46 | -6946.02 | 1.78 | 1.00 |
| DMA9_P6 | 0.51 | -6688.20 | 0.95 | 1.01 |
| DMA9_P7 | 0.81 | -6646.72 | -0.16 | 1.02 |
| DMA9_CP | 0.78 | -8809.38 | 1.34 | 1.03 |

# Appendix 2. Model-based localization results

## Localization error

| Date | DMA | Time [h] | Test Flow [l/s] | BASE | RE | MP | MP+RE |
|---|---|---|---|---|---|---|---|
| RUN1 | UK Winter time | | | | | | |
| 3-2-2022 | DMA2 | 09:00-10:30 | 0.8 | 0.09 | 0.09 | 0.41 | 0.41 |
| | | 10:30-11:30 | 1.5 | 0.09 | 0.09 | 0.41 | 0.41 |
| | DMA5 | 13:00-14:30 | 0.8 | 0.32 | 0.32 | 0.20 | 0.20 |
| | | 14:30-15:30 | 1.5 | 0.32 | 0.32 | 0.20 | 0.20 |
| 10-3-2022 | DMA8 | 09:00-10:30 | 0.8 | 0.61 | 0.61 | 0.00 | 0.19 |
| | | 10:30-11:30 | 1.5 | 0.60 | 0.55 | 0.61 | 0.17 |
| | DMA1 | 13:00-14:30 | 0.8 | 0.69 | 0.69 | 0.02 | 0.02 |
| | | 14:30-15:30 | 1.5 | 0.16 | 0.16 | 0.02 | 0.02 |
| 15-3-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.46 | 0.13 | 0.39 | 0.13 |
| | | 10:30-11:30 | 1.5 | 0.13 | 0.13 | 0.13 | 0.13 |
| | DMA3 | 13:00-14:30 | 0.8 | 0.30 | 0.30 | 0.19 | 0.03 |
| | | 14:30-15:30 | 1.5 | 0.30 | 0.30 | 0.03 | 0.03 |
| 23-3-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.98 | 0.72 | 0.98 | 0.81 |
| | | 10:30-11:30 | 1.5 | 0.98 | 0.77 | 0.98 | 0.72 |
| | DMA4 | 13:00-14:30 | 0.8 | 0.50 | 0.33 | 0.40 | 0.33 |
| | | 14:30-15:30 | 1.5 | 0.50 | 0.33 | 0.39 | 0.33 |
| 24-3-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.33 | 0.29 | 0.12 | 0.12 |
| | | 10:30-11:30 | 1.5 | 0.33 | 0.29 | 0.12 | 0.12 |
| RUN2 | UK Summer Time | | | | | | |
| 3-5-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.13 | 0.07 | 0.39 | 0.68 |
| | | 10:30-11:30 | 1.5 | 0.13 | 0.33 | 0.13 | 0.68 |
| 4-5-2022 | DMA6 | 01:00-02:30 | 0.8 | 0.13 | 0.26 | 0.13 | 0.04 |
| | | 02:30-03:30 | 1.5 | 0.13 | 0.26 | 0.13 | 0.46 |
| 5-5-2022 | DMA3 | 09:00-10:30 | 0.8 | 0.30 | 0.30 | 0.03 | 0.03 |
| | | 10:30-11:30 | 1.5 | 0.30 | 0.30 | 0.03 | 0.03 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 10-5-2022 | DMA1 | 09:00-10:30 | 0.8 | 0.69 | 0.69 | 0.00 | 0.00 |
| | | 10:30-11:30 | 1.5 | 0.69 | 0.69 | 0.00 | 0.00 |
| 13-5-2022 | DMA5 | 09:00-10:30 | 0.8 | 0.32 | 0.32 | 0.40 | 0.40 |
| | | 10:30-11:30 | 1.5 | 0.32 | 0.32 | 0.40 | 0.40 |
| 16-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.54 | 0.73 | 0.51 | 0.55 |
| | | 10:30-11:30 | 1.5 | 0.68 | 0.73 | 0.50 | 0.16 |
| 17-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.66 | 0.16 | 0.66 | 0.17 |
| | | 10:30-11:30 | 1.5 | 0.66 | 0.16 | 0.66 | 0.17 |
| 18-5-2022 | DMA3 | 11:10-12:10 | 0.8 | 0.28 | 0.69 | 0.06 | 0.06 |
| | | 12:10-13:10 | 1.5 | 0.30 | 0.69 | 0.06 | 0.06 |
| | | 13:10-14:10 | 2 | 0.69 | 0.69 | 0.06 | 0.06 |
| | | 14:10-15:10 | 2.5 | 0.69 | 0.69 | 0.06 | 0.06 |
| 23-5-2022 | DMA4 | 11:00-12:00 | 0.8 | 0.64 | 0.24 | 0.06 | 0.06 |
| | | 12:00-13:00 | 1.5 | 0.64 | 0.24 | 0.64 | 0.06 |
| | | 13:00-14:00 | 2 | 0.64 | 0.24 | 0.41 | 0.06 |
| | | 14:00-15:00 | 2.5 | 0.24 | 0.24 | 0.41 | 0.09 |
| 25-5-2022 | DMA4 | 09:00-15:00 | 0.8 | 0.60 | 0.26 | 0.26 | 0.26 |
| 26-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.67 | 0.49 | 0.32 | 0.09 |
| | | 10:30-11:30 | 1.5 | 0.67 | 0.49 | 0.09 | 0.09 |
| 27-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.06 | 0.71 | 0.00 | 0.00 |
| | | 10:30-11:30 | 1.5 | 0.06 | 0.71 | 0.12 | 0.12 |
| 1-6-2022 | DMA3 | 01:00-02:30 | 0.8 | 0.30 | 0.30 | 0.11 | 0.11 |
| | | 02:30-03:30 | 1.5 | 0.30 | 0.19 | 0.11 | 0.11 |

## False-positive path

| Date | DMA | Test Run Time | Test Flow [l/s] | BASE | RE | MP | MP+RE |
|---|---|---|---|---|---|---|---|
| RUN1 | UK Winter Time | | | | | | |
| 3-2-2022 | DMA2 | 09:00-10:30 | 0.8 | 0.94 | 0.96 | 0.26 | 0.26 |
| | | 10:30-11:30 | 1.5 | 0.94 | 0.96 | 0.38 | 0.38 |
| | DMA5 | 13:00-14:30 | 0.8 | 0.13 | 0.13 | 0.08 | 0.08 |
| | | 14:30-15:30 | 1.5 | 0.13 | 0.13 | 0.08 | 0.08 |
| 10-3-2022 | DMA8 | 09:00-10:30 | 0.8 | 0.33 | 0.33 | 0.01 | 0.01 |
| | | 10:30-11:30 | 1.5 | 0.26 | 0.57 | 0.01 | 0.01 |
| | DMA1 | 13:00-14:30 | 0.8 | 0.28 | 0.28 | 0.03 | 0.03 |
| | | 14:30-15:30 | 1.5 | 0.07 | 0.07 | 0.06 | 0.06 |
| 15-3-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.02 | 0.01 | 0.02 | 0.02 |
| | | 10:30-11:30 | 1.5 | 0.01 | 0.01 | 0.02 | 0.02 |
| | DMA3 | 13:00-14:30 | 0.8 | 0.23 | 0.22 | 0.01 | 0.01 |
| | | 14:30-15:30 | 1.5 | 0.22 | 0.20 | 0.01 | 0.01 |
| 23-3-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.75 | 0.19 | 0.84 | 0.25 |
| | | 10:30-11:30 | 1.5 | 0.80 | 0.19 | 0.82 | 0.25 |
| | DMA4 | 13:00-14:30 | 0.8 | 0.56 | 0.45 | 0.45 | 0.39 |
| | | 14:30-15:30 | 1.5 | 0.54 | 0.36 | 0.45 | 0.39 |
| 24-3-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.12 | 0.30 | 0.11 | 0.09 |
| | | 10:30-11:30 | 1.5 | 0.31 | 0.30 | 0.09 | 0.08 |
| RUN2 | UK Summer Time | | | | | | |
| 3-5-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.01 | 0.79 | 0.01 | 0.05 |
| | | 10:30-11:30 | 1.5 | 0.01 | 0.84 | 0.01 | 0.10 |
| 4-5-2022 | DMA6 | 01:00-02:30 | 0.8 | 0.01 | 0.92 | 0.01 | 0.24 |
| | | 02:30-03:30 | 1.5 | 0.01 | 0.92 | 0.01 | 0.24 |
| 5-5-2022 | DMA3 | 09:00-10:30 | 0.8 | 0.17 | 0.17 | 0.03 | 0.03 |
| | | 10:30-11:30 | 1.5 | 0.22 | 0.22 | 0.01 | 0.01 |
| 10-5-2022 | DMA1 | 09:00-10:30 | 0.8 | 0.30 | 0.30 | 0.00 | 0.00 |
| | | 10:30-11:30 | 1.5 | 0.28 | 0.28 | 0.00 | 0.00 |
| 13-5-2022 | DMA5 | 09:00-10:30 | 0.8 | 0.13 | 0.13 | 0.78 | 0.78 |
| | | 10:30-11:30 | 1.5 | 0.12 | 0.12 | 0.87 | 0.87 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 16-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.31 | 0.34 | 0.30 | 0.06 |
| | | 10:30-11:30 | 1.5 | 0.35 | 0.34 | 0.34 | 0.07 |
| 17-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.57 | 0.29 | 0.32 | 0.05 |
| | | 10:30-11:30 | 1.5 | 0.57 | 0.29 | 0.52 | 0.02 |
| 18-5-2022 | DMA3 | 11:10-12:10 | 0.8 | 0.22 | 0.22 | 0.16 | 0.12 |
| | | 12:10-13:10 | 1.5 | 0.23 | 0.24 | 0.17 | 0.16 |
| | | 13:10-14:10 | 2 | 0.43 | 0.44 | 0.07 | 0.07 |
| | | 14:10-15:10 | 2.5 | 0.32 | 0.51 | 0.07 | 0.07 |
| 23-5-2022 | DMA4 | 11:00-12:00 | 0.8 | 0.35 | 0.92 | 0.04 | 0.04 |
| | | 12:00-13:00 | 1.5 | 0.33 | 0.92 | 0.24 | 0.03 |
| | | 13:00-14:00 | 2 | 0.33 | 0.92 | 0.10 | 0.03 |
| | | 14:00-15:00 | 2.5 | 0.34 | 0.89 | 0.11 | 0.10 |
| 25-5-2022 | DMA4 | 09:00-15:00 | 0.8 | 0.54 | 0.11 | 0.02 | 0.10 |
| 26-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.28 | 0.27 | 0.13 | 0.12 |
| | | 10:30-11:30 | 1.5 | 0.31 | 0.28 | 0.13 | 0.12 |
| 27-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.99 | 0.74 | 0.00 | 0.00 |
| | | 10:30-11:30 | 1.5 | 0.99 | 0.70 | 0.00 | 0.00 |
| 1-6-2022 | DMA3 | 01:00-02:30 | 0.8 | 0.34 | 0.52 | 0.07 | 0.07 |
| | | 02:30-03:30 | 1.5 | 0.28 | 0.44 | 0.07 | 0.07 |

# Appendix 3. Model-based leak localization with interpolation

## Localization error

| Date | DMA | Time [h] | Test Flow [l/s] | BASE+GP | RE+GP | MP+GP | MP+RE+GP |
|---|---|---|---|---|---|---|---|
| RUN1 | UK Winter time | | | | | | |
| 3-2-2022 | DMA2 | 09:00-10:30 | 0.8 | 0.42 | 0.42 | 0.41 | 0.41 |
| | | 10:30-11:30 | 1.5 | 0.33 | 0.33 | 0.41 | 0.41 |
| | DMA5 | 13:00-14:30 | 0.8 | 0.35 | 0.35 | 0.33 | 0.33 |
| | | 14:30-15:30 | 1.5 | 0.36 | 0.36 | 0.33 | 0.33 |
| 10-3-2022 | DMA8 | 09:00-10:30 | 0.8 | 0.36 | 0.39 | 0.58 | 0.58 |
| | | 10:30-11:30 | 1.5 | 0.59 | 0.37 | 0.58 | 0.60 |
| | DMA1 | 13:00-14:30 | 0.8 | 0.00 | 0.00 | 0.12 | 0.12 |
| | | 14:30-15:30 | 1.5 | 0.00 | 0.00 | 0.12 | 0.12 |
| 15-3-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.39 | 0.67 | 0.64 | 0.39 |
| | | 10:30-11:30 | 1.5 | 0.39 | 0.67 | 0.64 | 0.20 |
| | DMA3 | 13:00-14:30 | 0.8 | 0.30 | 0.88 | 0.15 | 0.16 |
| | | 14:30-15:30 | 1.5 | 0.19 | 0.88 | 0.15 | 0.16 |
| 23-3-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.98 | 0.18 | 0.90 | 0.79 |
| | | 10:30-11:30 | 1.5 | 0.98 | 0.16 | 0.89 | 0.85 |
| | DMA4 | 13:00-14:30 | 0.8 | 0.52 | 0.42 | 0.33 | 0.33 |
| | | 14:30-15:30 | 1.5 | 0.52 | 0.41 | 0.33 | 0.33 |
| 24-3-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.48 | 0.29 | 0.26 | 0.26 |
| | | 10:30-11:30 | 1.5 | 0.29 | 0.29 | 0.26 | 0.26 |
| RUN2 | UK Summer Time | | | | | | |
| 3-5-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.39 | 0.67 | 0.64 | 0.67 |
| | | 10:30-11:30 | 1.5 | 0.13 | 0.67 | 0.64 | 0.67 |
| 4-5-2022 | DMA6 | 01:00-02:30 | 0.8 | 0.67 | 0.67 | 0.64 | 0.67 |
| | | 02:30-03:30 | 1.5 | 0.13 | 0.67 | 0.64 | 0.67 |
| 5-5-2022 | DMA3 | 09:00-10:30 | 0.8 | 0.45 | 0.45 | 0.15 | 0.15 |
| | | 10:30-11:30 | 1.5 | 0.36 | 0.36 | 0.30 | 0.30 |
| 10-5-2022 | DMA1 | 09:00-10:30 | 0.8 | 0.69 | 0.69 | 0.02 | 0.02 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | 10:30-11:30 | 1.5 | 0.69 | 0.69 | 0.02 | 0.02 |
| 13-5-2022 | DMA5 | 09:00-10:30 | 0.8 | 0.33 | 0.33 | 0.33 | 0.33 |
| | | 10:30-11:30 | 1.5 | 0.33 | 0.33 | 0.33 | 0.33 |
| 16-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.63 | 0.63 | 0.31 | 0.55 |
| | | 10:30-11:30 | 1.5 | 0.48 | 0.48 | 0.31 | 0.44 |
| 17-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.68 | 0.68 | 0.52 | 0.42 |
| | | 10:30-11:30 | 1.5 | 0.68 | 0.68 | 0.50 | 0.50 |
| 18-5-2022 | DMA3 | 11:10-12:10 | 0.8 | 0.06 | 0.68 | 0.74 | 0.74 |
| | | 12:10-13:10 | 1.5 | 0.00 | 0.74 | 0.71 | 0.74 |
| | | 13:10-14:10 | 2 | 0.28 | 0.74 | 0.71 | 0.74 |
| | | 14:10-15:10 | 2.5 | 0.00 | 0.74 | 0.74 | 0.74 |
| 23-5-2022 | DMA4 | 11:00-12:00 | 0.8 | 0.07 | 0.89 | 0.72 | 0.72 |
| | | 12:00-13:00 | 1.5 | 0.07 | 0.72 | 0.07 | 0.68 |
| | | 13:00-14:00 | 2 | 0.07 | 0.72 | 0.07 | 0.68 |
| | | 14:00-15:00 | 2.5 | 0.07 | 0.72 | 0.07 | 0.68 |
| 25-5-2022 | DMA4 | 09:00-15:00 | 0.8 | 0.63 | 0.29 | 0.26 | 0.26 |
| 26-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.22 | 0.17 | 0.65 | 0.65 |
| | | 10:30-11:30 | 1.5 | 0.17 | 0.34 | 0.65 | 0.65 |
| 27-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.71 | 0.75 | 0.05 | 0.13 |
| | | 10:30-11:30 | 1.5 | 0.71 | 0.71 | 0.05 | 0.13 |
| 1-6-2022 | DMA3 | 01:00-02:30 | 0.8 | 0.11 | 0.88 | 0.02 | 0.00 |
| | | 02:30-03:30 | 1.5 | 0.02 | 0.62 | 0.00 | 0.00 |

## False-positive path

| Date | DMA | Time [h] | Test Flow [l/s] | BASE+GP | RE+GP | MP+GP | MP+RE+GP |
|---|---|---|---|---|---|---|---|
| RUN1 | UK Winter Time | | | | | | |
| 3-2-2022 | DMA2 | 09:00-10:30 | 0.8 | 0.24 | 0.24 | 0.67 | 0.67 |
| | | 10:30-11:30 | 1.5 | 0.28 | 0.28 | 0.67 | 0.67 |
| | DMA5 | 13:00-14:30 | 0.8 | 0.45 | 0.45 | 0.33 | 0.33 |
| | | 14:30-15:30 | 1.5 | 0.60 | 0.60 | 0.32 | 0.32 |
| 10-3-2022 | DMA8 | 09:00-10:30 | 0.8 | 0.59 | 0.42 | 0.54 | 0.52 |
| | | 10:30-11:30 | 1.5 | 0.61 | 0.43 | 0.52 | 0.50 |
| | DMA1 | 13:00-14:30 | 0.8 | 0.00 | 0.00 | 0.14 | 0.14 |
| | | 14:30-15:30 | 1.5 | 0.00 | 0.00 | 0.14 | 0.14 |
| 15-3-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.28 | 0.51 | 0.39 | 0.46 |
| | | 10:30-11:30 | 1.5 | 0.29 | 0.51 | 0.38 | 0.41 |
| | DMA3 | 13:00-14:30 | 0.8 | 0.30 | 0.19 | 0.21 | 0.07 |
| | | 14:30-15:30 | 1.5 | 0.30 | 0.14 | 0.25 | 0.10 |
| 23-3-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.83 | 0.05 | 0.24 | 0.21 |
| | | 10:30-11:30 | 1.5 | 0.84 | 0.05 | 0.24 | 0.20 |
| | DMA4 | 13:00-14:30 | 0.8 | 0.54 | 0.46 | 0.57 | 0.60 |
| | | 14:30-15:30 | 1.5 | 0.54 | 0.46 | 0.50 | 0.46 |
| 24-3-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.33 | 0.33 | 0.65 | 0.65 |
| | | 10:30-11:30 | 1.5 | 0.33 | 0.25 | 0.64 | 0.64 |
| RUN2 | UK Summer Time | | | | | | |
| 3-5-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.18 | 0.51 | 0.39 | 0.43 |
| | | 10:30-11:30 | 1.5 | 0.05 | 0.51 | 0.38 | 0.50 |
| 4-5-2022 | DMA6 | 01:00-02:30 | 0.8 | 0.37 | 0.53 | 0.38 | 0.54 |
| | | 02:30-03:30 | 1.5 | 0.01 | 0.57 | 0.38 | 0.54 |
| 5-5-2022 | DMA3 | 09:00-10:30 | 0.8 | 0.39 | 0.39 | 0.19 | 0.19 |
| | | 10:30-11:30 | 1.5 | 0.39 | 0.39 | 0.30 | 0.30 |
| 10-5-2022 | DMA1 | 09:00-10:30 | 0.8 | 0.31 | 0.31 | 0.03 | 0.03 |
| | | 10:30-11:30 | 1.5 | 0.31 | 0.31 | 0.14 | 0.14 |
| 13-5-2022 | DMA5 | 09:00-10:30 | 0.8 | 0.49 | 0.49 | 0.63 | 0.63 |
| | | 10:30-11:30 | 1.5 | 0.35 | 0.35 | 0.62 | 0.62 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 16-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.31 | 0.31 | 0.45 | 0.69 |
| | | 10:30-11:30 | 1.5 | 0.31 | 0.30 | 0.44 | 0.69 |
| 17-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.49 | 0.61 | 0.60 | 0.35 |
| | | 10:30-11:30 | 1.5 | 0.55 | 0.6 | 0.63 | 0.35 |
| 18-5-2022 | DMA3 | 11:10-12:10 | 0.8 | 0.07 | 0.18 | 0.01 | 0.00 |
| | | 12:10-13:10 | 1.5 | 0.01 | 0.19 | 0.00 | 0.00 |
| | | 13:10-14:10 | 2 | 0.07 | 0.60 | 0.50 | 0.53 |
| | | 14:10-15:10 | 2.5 | 0.00 | 0.57 | 0.45 | 0.52 |
| 23-5-2022 | DMA4 | 11:00-12:00 | 0.8 | 0.08 | 0.91 | 0.90 | 0.90 |
| | | 12:00-13:00 | 1.5 | 0.08 | 0.91 | 0.08 | 0.73 |
| | | 13:00-14:00 | 2 | 0.08 | 0.90 | 0.08 | 0.42 |
| | | 14:00-15:00 | 2.5 | 0.08 | 0.90 | 0.08 | 0.42 |
| 25-5-2022 | DMA4 | 09:00-15:00 | 0.8 | 0.61 | 0.38 | 0.30 | 0.46 |
| 26-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.07 | 0.04 | 0.78 | 0.78 |
| | | 10:30-11:30 | 1.5 | 0.07 | 0.06 | 0.78 | 0.78 |
| 27-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.97 | 0.69 | 0.41 | 0.42 |
| | | 10:30-11:30 | 1.5 | 0.97 | 0.67 | 0.41 | 0.42 |
| 1-6-2022 | DMA3 | 01:00-02:30 | 0.8 | 0.18 | 0.53 | 0.45 | 0.53 |
| | | 02:30-03:30 | 1.5 | 0.00 | 0.50 | 0.45 | 0.49 |

# Appendix4. Data-driven leak localization results

| Date | DMA | Time [h] | Test Flow [l/s] | LE | Pipe length [m] | Distance [m] |
|---|---|---|---|---|---|---|
| RUN1 | UK Winter Time | | | | | |
| 3-2-2022 | DMA5 | 13:00-14:30 | 0.8 | 0.20 | 210 | 189 |
| | | 14:30-15:30 | 1.5 | 0.20 | 210 | 189 |
| 10-3-2022 | DMA1 | 13:00-14:30 | 0.8 | 0.12 | 356 | 339 |
| | | 14:30-15:30 | 1.5 | 0.12 | 356 | 339 |
| 15-3-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.12 | 187 | 171 |
| | | 10:30-11:30 | 1.5 | 0.12 | 187 | 171 |
| | DMA3 | 13:00-14:30 | 0.8 | 0.02 | 99 | 67 |
| | | 14:30-15:30 | 1.5 | 0.02 | 99 | 67 |
| 23-3-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.70 | 1114 | 634 |
| | | 10:30-11:30 | 1.5 | 0.70 | 1114 | 634 |
| | DMA4 | 13:00-14:30 | 0.8 | 0.07 | 74 | 64 |
| | | 14:30-15:30 | 1.5 | 0.07 | 74 | 64 |
| 24-3-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.49 | 709 | 509 |
| | | 10:30-11:30 | 1.5 | 0.49 | 709 | 509 |
| RUN2 | UK Summer Time | | | | | |
| 3-5-2022 | DMA6 | 09:00-10:30 | 0.8 | 0.12 | 187 | 171 |
| | | 10:30-11:30 | 1.5 | 0.12 | 187 | 171 |
| 4-5-2022 | DMA6 | 01:00-02:30 | 0.8 | 0.24 | 501 | 417 |
| | | 02:30-03:30 | 1.5 | 0.24 | 501 | 417 |
| 5-5-2022 | DMA3 | 09:00-10:30 | 0.8 | 0.02 | 99 | 67 |
| | | 10:30-11:30 | 1.5 | 0.02 | 99 | 67 |
| 10-5-2022 | DMA1 | 09:00-10:30 | 0.8 | 0.12 | 356 | 339 |
| | | 10:30-11:30 | 1.5 | 0.12 | 356 | 339 |
| 13-5-2022 | DMA5 | 09:00-10:30 | 0.8 | 0.33 | 761 | 578 |
| | | 10:30-11:30 | 1.5 | 0.33 | 761 | 578 |
| 16-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.47 | 812 | 569 |
| | | 10:30-11:30 | 1.5 | 0.47 | 812 | 569 |
| 17-5-2022 | DMA10 | 09:00-10:30 | 0.8 | 0.17 | 478 | 279 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | 10:30-11:30 | 1.5 | 0.17 | 478 | 279 |
| 18-5-2022 | DMA3 | 11:10-12:10 | 0.8 | 0.30 | 540 | 377 |
| | | 12:10-13:10 | 1.5 | 0.30 | 540 | 377 |
| | | 13:10-14:10 | 2 | 0.30 | 540 | 377 |
| | | 14:10-15:10 | 2.5 | 0.30 | 540 | 377 |
| 23-5-2022 | DMA4 | 11:00-12:00 | 0.8 | 0.07 | 276 | 81 |
| | | 12:00-13:00 | 1.5 | 0.07 | 276 | 81 |
| | | 13:00-14:00 | 2 | 0.07 | 276 | 81 |
| | | 14:00-15:00 | 2.5 | 0.07 | 276 | 81 |
| 25-5-2022 | DMA4 | 09:00-15:00 | 0.8 | 0.00 | 0 | 5 |
| 26-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.22 | 380 | 263 |
| | | 10:30-11:30 | 1.5 | 0.22 | 380 | 263 |
| 27-5-2022 | DMA7 | 09:00-10:30 | 0.8 | 0.74 | 1079 | 741 |
| | | 10:30-11:30 | 1.5 | 0.74 | 1079 | 741 |
| 1-6-2022 | DMA3 | 01:00-02:30 | 0.8 | 0.02 | 99 | 67 |
| | | 02:30-03:30 | 1.5 | 0.02 | 99 | 67 |