Leakage diagnosis with a contamination mitigation control framework using a graph theory based model

Rathore, Saruch Satishkumar; Misra, Rahul; Kallesøe, Carsten Skovmose; Wisniewski, Rafal

Published in:
Annual Reviews in Control

DOI (link to publication from Publisher):
10.1016/j.arcontrol.2023.03.010

Publication date:
2023

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

Annual Reviews in Control 55 (2023) 498–519

Contents lists available at ScienceDirect

Annual Reviews in Control

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

Full Length Article

Leakage diagnosis with a contamination mitigation control framework using a graph theory based model ☆

Saruch Satishkumar Rathore [a,*], Rahul Misra [a], Carsten Skovmose Kallesøe [a,b], Rafal Wisniewski [a]

[a] Aalborg University, Fredrik Bajers Vej 7, Aalborg, DK-9220, Denmark
[b] Grundfos Holding A/S, Poul Due Jensens Vej 7, Bjerringbro, DK-8850, Denmark

ARTICLE INFO

Keywords:
Water distribution networks
Graph theory
Leakage diagnosis
Contamination mitigation control
Water contamination

ABSTRACT

This work proposes an operational management approach for water distribution networks (WDNs) that can detect and localize leakages while also mitigating contamination resulting from these leaks. The primary emphasis of this work is the development of a contamination mitigation control scheme. A leak typically leads to a drop in network pressure that increases the risk of contamination. A leakage localization algorithm is responsible for detecting and localizing the leakage in the WDN. When a leak is detected in the network the contamination mitigation control is activated. The flow and pressure settings of the pumps are regulated by the contamination mitigation control in an optimal manner to minimize the risk of contamination. The entire framework is tested on the Smart Water Infrastructure Laboratory situated at Aalborg University, Denmark and a large-scale benchmark water network, which is part of a city network, L-town.

1. Introduction

Water scarcity is a global issue that affects millions of people worldwide. According to UN-Water (2021), approximately 2.3 billion people, or about one-third of the global population, reside in water-stressed countries. Of these, 733 million live in countries experiencing high or critically high levels of water stress. These numbers are expected to worsen due to the increasing demand for water and climate change. In spite of that, a significant amount of water is lost due to leakages in the water distribution networks. The volume of non-revenue water is estimated to be 126 billion cubic metres (Liemberger & Wyatt, 2019). Apart from water loss, leakages also have negative economic and social impacts with damage to infrastructure and disruption to the community.

Leakages in a water distribution network can not only result in wasted resources but also impact the quality of the water. When leakages occur in the pipelines, it can lead to the contamination of the water with various substances, including chemicals, bacteria, and other pollutants (Ebacher, Besner, Prévost, & Allard, 2010). According to experimental test results presented by Fontanazza, Notaro, Puleo, Nicolosi, and Freni (2015), cracks in pipelines can allow contaminants to enter the water distribution network due to low or negative pressures. Similar findings have been reported in other studies that investigated contaminant intrusion through leakage points during negative pressure events, such as Besner et al. (2010), Besner, Prévost, and Regli (2011), Karim, Abbaszadegan, and LeChevallier (2003) and Yang, LeChevallier, Teunis, and Xu (2011). Water contamination can present significant risks to public health (Islam, Farahat, Al-Zahrani, Rodriguez, & Sadiq, 2015). Kirmeyer (2001) documents numerous instances of disease outbreaks resulting from waterborne pollutants entering through pipe ruptures and it classifies contamination occurring due to water main break/repair as high risk. Around 9% water contamination regarded illness outbreak occurred due to pipe breaks in the water networks in the USA from 1971–1998 (Sadiq, Kleiner, & Rajani, 2006). Further, even in case of leakages, the water utilities need to ensure the safety and reliability of the water supply to the consumers. These challenges highlight the importance of effective leakage diagnosis and water management.

Water distribution networks (WDN) usually have a complex topology making the leakage diagnosis process troublesome. Several water utilities rely on techniques such as Minimum Night Flow (MNF) (Mazzolani et al., 2017), for leakage detection and identification. Other commercially available methods include leakage detection using electronic amplified listening devices and acoustic noise loggers (Hamilton & Charalambous, 2013). These traditional methods either requires expensive equipment or manual labour, are highly human-skill dependent, have low accuracy and are time-consuming. Also, with new infrastructures with larger diameter pipes or plastic pipes, these methods have proven to be less effective (Hamilton & Charalambous, 2013).

☆ This work was supported by Poul Due Jensen Foundation, Denmark.
* Corresponding author.
E-mail addresses: ssr@es.aau.dk (S.S. Rathore), rmi@es.aau.dk (R. Misra), csk@es.aau.dk (C.S. Kallesøe), raf@es.aau.dk (R. Wisniewski).

https://doi.org/10.1016/j.arcontrol.2023.03.010
Received 23 December 2022; Accepted 21 March 2023
Available online 12 April 2023

The water sector is moving towards digitalization and water utilities around the world are installing sensors in the network. There are many works focusing on leakage diagnosis using these sensors. Adedeji, Hamam, Abe, and Abu-Mahfouz (2017), Chan, Chin, and Zhong (2018) and Puust, Kapelan, Savic, and Koppel (2010) provides a comprehensive review and comparison of such work for leakage detection and isolation in water distribution networks is presented. This work particularly focuses on leakage detection and isolation using pressure sensors. Leakages in a water distribution network, notably burst leakages, often result in sudden unexpected deviations in pressure within the network. By analysing this pressure data, it may be possible to identify and diagnose the leakage. Various emerging techniques for leakage diagnosis based on pressure data are presented and compared, on benchmark data, by Vrachimis et al. (2022). Geometric leakage isolation has been considered by Casillas, Garza-Castañón, Puig, and Vargas-Martinez (2015), where the localization of the leakages is based on residual vectors, which are compared with pre-calculated signature vectors. Leakage localization using fault sensitivity matrices was explored by Pérez et al. (2011) and Perez et al. (2014), where the sensitivity of pressure measurements to fault scenarios was used and the residual vectors are compared to the sensitivity matrix. Various methods for this comparison are compared and presented by Casillas, Garza-Castañón, and Puig (2013), Casillas Ponce, Garza Castanon, and Cayuela (2014) and Ponce, Castañón, and Cayuela (2012). However, in all these works the sensitivity matrix was computed using a standard hydraulic simulation model. Whereas, Rathore, Kallesøe, and Wisniewski (2022) and Rathore, Kallesøe, Wisniewski, and Jensen (2021) derive the sensitivity matrix using a reduced order model and consider leakages as a change in the distribution among the consumers rather than a change in consumption.

The problem of water contamination has been addressed in numerous studies, with contamination detection being the first step. Lambrou, Anastasiou, Panayiotou, and Polycarpou (2014) proposed a novel approach for detecting contamination using low-cost sensors. In addition, model-based methods for contamination event detection using Monte-Carlo simulations have been presented by Eliades, Lambrou, Panayiotou, and Polycarpou (2014) and Eliades, Stavrou, Vrachimis, Panayiotou, and Polycarpou (2015). Forecasting the spread of contaminants in the network, viz. contamination zoning map, using rough set theory is presented by Bazargan-Lari, Taghipour, and Habibi (2021). After contamination has been detected, mitigation strategies for addressing water contamination may include response actions such as warning consumers to reduce water usage and flushing contaminated water through hydrants (Shafiee & Berglund, 2017). Di Nardo, Di Natale, Guida, and Musmarra (2013) studies the effect of sectionalizing the water network into district meter areas and district isolation to reduce the impact of contamination. The design of the actuator network in water networks to isolate the contamination is presented by Palleti, Kurian, Narasimhan, and Rengaswamy (2018). A heuristic approach is presented by Poulin et al. (2008), where the network valves are closed to isolate the contamination and then the contamination is removed by flushing. Numerous studies have formulated the contaminant flushing problem as a single or multiple objective optimization problem. The operation of the actuators, such as valves, hydrants and/or pumps in the network, is defined as the decision variables in this optimization problem. The objective function in this problem is a combination of the following objectives: (a) minimizing the health impact on the consumers, (b) minimizing the disruption of service (c) minimizing the number of operations. Then the optimization problem is solved using various methods such as genetic algorithm (Rasekh & Brumbelow, 2014), deep reinforcement learning (Hu, Wang, Gong, & Yan, 2022), evolutionary optimization (Alfonso, Jonoski, & Solomatine, 2010) and swarm optimization (Moghaddam, Afsharnia, & Peirovi Minaee, 2020). Shafiee and Berglund (2015) presents a decision tree-based approach for contaminant flushing where the decision tree is obtained offline using a noisy genetic algorithm. NSGA-II algorithm coupled with the hydraulic simulator, EPANET (Rossman et al., 2000), has been used to simulate multiple scenarios of contamination and compute response objective values by Guidorzi, Franchini, and Alvisi (2009) and Preis and Ostfeld (2008a). Overall, these studies provide valuable insights and solutions for detecting and addressing water contamination.

In this work, we extend the work of Rathore et al. (2022) and Rathore et al. (2021) on leakage detection and localization to add contamination mitigation control. As presented by Rathore et al. (2022), the leakage localization is based on pressure residuals generated using a data-driven model. Then the pressure residuals are compared to the sensitivity matrix to identify the probable location of the leakage. The pressure within the network usually reacts differently depending on the location of the leak. This method involves comparing the pressure residual to the pressure sensitivity for leaks at different locations to estimate the actual location of the leak. Since this approach is based on pressure sensor data, which is relatively inexpensive, it is economically viable. Furthermore, the sensitivity matrix and leakage location indicators are obtained through matrix computation, so the method is not computationally intensive. Based on the information about the leakage, a contamination mitigation control is designed in this work. The previous works on contamination response focused on moving the contaminant from a source location to a hydrant or sink and then flushing it out of the network from there. However, this work focuses only on contamination caused by leakages, with the objective of preventing the entry of the contaminant from the leakage point or limiting the spread of contamination to the area of the leakage. The idea is to maintain the water flow in the network towards the leakage area. This can be viewed as flushing from the leakage point until the leakage is repaired. This would prevent water contamination from happening in the first place. A novel method for identifying the pipes and their flow direction is also presented, which allows for the water to be directed towards the leakage area. While the water is being flushed to prevent contamination, consumer demands are also to be fulfilled. Therefore, the contamination mitigation control is formulated as an optimization problem, in which consumer demands are also taken into account. In this work, a mathematical model of water distribution networks based on graph theory serves as the foundation for both the leakage diagnosis algorithm and contamination mitigation control. This model, first introduced by Kallesøe, Jensen, and Wisniewski (2015), represents the network in a steady state with one pressure and one flow equation, making it well-suited for developing a variety of control schemes and applications. The mathematical model is constructed with the graph incidence matrix of the network and network parameters such as pipe length, diameter and elevation. The network matrix and parameters can be easily obtained from EPANET models of the network using various programming toolboxes such as Eliades, Kyriakou, Vrachimis, and Polycarpou (2016). Previous works on leakage diagnosis based sensitivity analysis, as mentioned earlier, use an EPANET model to compute the sensitivity matrix by running multiple simulations of different leakage scenarios which could be a cumbersome task for large networks. However, in this work and in Rathore et al. (2022), we use the leakage diagnosis algorithm developed upon this graphic theory based model. The sensitivity matrix is also derived instead of being computed from simulations. The contamination mitigation control, which is the main contribution of this paper, is also developed upon this model. As previously mentioned, the control is formulated as an optimization problem and the network model equations are incorporated into the optimization problem as constraints. In contrast, previous works on contamination flushing, such as Preis and Ostfeld (2008a) and Rasekh and Brumbelow (2014), relied on EPANET simulations for optimal strategy, which again can be a cumbersome task. Finally, the leakage diagnosis with contamination mitigation control framework is tested on a laboratory water network setup and the large-scale benchmark water network model in EPANET.

The rest of the paper is organized as follows. Section 2 presents some of the common notations used in the work. As a preliminary

to this work, a graph theory based mathematical model for a water distribution network is presented in Section 3. The graph based modelling framework creates the fundamentals for generalizing the results to any network without elevated reservoirs. In Section 4, the leakage diagnosis methodology is presented. Section 5 presents the design of the contamination mitigation control. In Section 6, the results from the laboratory test and the test on the benchmark water network are presented. The limitations of the work and probable future works are discussed in Section 7. Finally, Section 8 concludes the work.

2. Notations

This section aims to put forward the common notations used throughout the paper. For a vector x, $x \in \mathbb{R}^n_{\geqslant 0}$ denotes $\forall i\ x_i \geq 0$, similarly $x \in \mathbb{R}^n_{\leqslant 0}$ denotes $\forall i\ x_i \leq 0$. The vector $\mathbb{1}$ is a column vector of ones in all positions and of appropriate dimension. The notation $\partial_x f|_{X,Y}$ is used to denote the partial derivative of f with respect to x evaluated at X, Y. For any two vectors x and y, $\langle x, y\rangle$ denotes the inner product between x and y. For a vector x, the notation $diag(x)$ is used to denote a diagonal matrix with elements of x on the matrix diagonal. The notation $E^{|\cdot|}$ is used to denote the element-wise absolute value of a matrix E.

The model used in the derivation of the leakage localization scheme and for control of pumping stations is derived using graph theory. We denote the variable corresponding to the spanning tree of the graph by the subscript $\mathcal{T}$ and the corresponding chord by the subscript $\mathcal{C}$. The model variables will be updated at discrete time instants and these time instants will be denoted by $[k]$. Furthermore, any matrix or vector with a row corresponding to the reference node removed is denoted with a bar, for example, $\bar{H}$, $\bar{d}$.

3. Preliminary on graph theory based mathematical model of a water distribution network

In this section, we recall the graph theory based mathematical model of a water distribution network which was first presented by Kallesøe et al. (2015). In Section 5, based on this model the contamination mitigation control strategy is developed. Further, under some assumptions, a reduced order model is presented, on which the leakage diagnosis algorithm is presented in Section 4.

A water distribution network can be represented as a directed graph (digraph), $\mathcal{G}$ (Deo, 2017). Then the m pipes of the network are represented by the m edges of the graph and the n connection points are represented by the n nodes of the graph. In this model, the supply flows and the consumer demands are modelled by assigning independent nodal flows to a subset of the nodes. Further, the graph $\mathcal{G}$ can be represented by an incidence matrix $H = [h_{ij}]$, where the element h_{ij} is defined as (Deo, 2017)

$$h_{ij} = \begin{cases} -1 & \text{if the } j\text{th edge is entering } i\text{th node.} \\ 0 & \text{if the } j\text{th edge is not connected to the } i\text{th node.} \\ 1 & \text{if the } j\text{th edge is leaving } i\text{th node.} \end{cases} \tag{1}$$

The direction of the edge is used to track the direction of the flow in the corresponding pipe. With this definition, each row of $H \in \{-1, 0, 1\}^{n\times m}$ corresponds to a node and each column corresponds to an edge in $\mathcal{G}$. $(n-1)$ rows of H contains all the information of H and therefore any one row can be removed, (Deo, 2017). The matrix obtained by removing an arbitrary row from H is referred to as a reduced incidence matrix, $\bar{H} \in \{-1, 0, 1\}^{n-1\times m}$, and the node corresponding to the removed row is refereed as the reference node. In this work, without loss of generality, we set the nth node as the reference node. Furthermore, the graph $\mathcal{G}$ can be divided into an arbitrary spanning tree $\mathcal{T}$ and its corresponding chords $\mathcal{C}$. Likewise $\bar{H}$ can also be partitioned into sub-matrices $\bar{H}_\mathcal{T}$ and $\bar{H}_\mathcal{C}$.

$$\bar{H} = \begin{bmatrix} \bar{H}_\mathcal{C} & \bar{H}_\mathcal{T} \end{bmatrix} \tag{2}$$

Now, for a water distribution network, Kirchhoff's node law can be given as,

$$Hq[k] = d[k], \tag{3}$$

where $q[k] \in \mathbb{R}^m$ is the vector of flows through the edges and $d[k] \in \mathbb{R}^n$ is the vector of independent nodal flows at each node which, in this work, is used for modelling consumer demands and supply flow. Note that, the nodal flows are modelled as independent flows of water in or out of the network nodes, i.e. the inlet flows and the consumer demands are not considered to be pressure dependent. It is assumed that a minimum pressure is always maintained at the nodes to satisfy consumer demands. We assume that users either manually adjust the flow according to pressure changes (e.g. when taking a shower) or use a given volume (e.g. when filling a bucket), both of which result in a controlled average flow over a specified time interval. For non-supply nodes where the flow is out of the network, $d_i[k] \leqslant 0$ and for the supply nodes, where the flow is into the network, $d_i[k] \geqslant 0$.

Also, the vector of flows through edges can be partitioned as $q^T = [q_\mathcal{C}^T\ q_\mathcal{T}^T]^T$, where $q_\mathcal{T}$ is the vector of flows through the spanning tree and $q_\mathcal{C}$ is the vector of flows through the corresponding chords. With (2), $q_\mathcal{T}[k]$ can be found by solving the part of (3) which is related to the non-reference nodes,

$$q_\mathcal{T}[k] = -\bar{H}_\mathcal{T}^{-1}\bar{H}_\mathcal{C} q_\mathcal{C}[k] + \bar{H}_\mathcal{T}^{-1}\bar{d}[k], \tag{4}$$

where $\bar{d}[k] \in \mathbb{R}^{(n-1)}$ the vector of nodal flows of the non-reference nodes.

Let p be the vector of absolute pressures at the nodes and Δp be the vector of the differential pressure across the edges, i.e. for an edge l connecting the nodes i with j, $\Delta p_l[k] = p_j[k] - p_i[k]$, where i is the entering node and j the leaving node of the edge. With this definition, the pressure drop over the edges can be given as,

$$\Delta p[k] = \lambda(q[k]) - \Delta z = H^T p[k], \tag{5}$$

where $\lambda(q[k]) = \begin{bmatrix}\lambda_1(q_1[k]) & \cdots & \lambda_m(q_m[k]))\end{bmatrix}^T$ models the flow-dependent pressure drops and $\lambda_i : \mathbb{R} \to \mathbb{R}$ is assumed to be of the form,

$$\lambda_i(q_i) = f_i|q_i|q_i \quad \text{with} \quad f_i > 0. \tag{6}$$

This expression of λ_i is an approximation for a turbulent flow, which is typical for a water supply application (Swamee & Sharma, 2008). The component $\Delta z \in \mathbb{R}^m$ is the vector of pressure drops due to geodesic level differences between nodes. Also, from graph theory $\Delta z = H^T z$, where $z \in \mathbb{R}^n$ is a vector of pressures due to geodesic levels at the nodes.

Furthermore, let B be the fundamental cycle matrix of $\mathcal{G}$ with respect to $\mathcal{T}$. A construction of B is (Deo, 2017),

$$B = \begin{bmatrix} I & -\bar{H}_\mathcal{C}^T \bar{H}_\mathcal{T}^{-T}. \end{bmatrix} \tag{7}$$

Multiplying (5) by B on the left side leads to $B\lambda(q[k]) = 0$, by the virtue of Kirchhoff's mesh law for hydraulic networks which gives $B\Delta z = 0$ and $B(\Delta p[k] + \Delta z) = 0$. Again partitioning $\lambda(q[k])^T = [\lambda(q[k])_\mathcal{C}^T\ \lambda(q[k])_\mathcal{T}^T]^T$ and using (4) gives implicit expression for the chord flows as,

$$\lambda_\mathcal{C}(q_\mathcal{C}[k]) - \bar{H}_\mathcal{C}^T \bar{H}_\mathcal{T}^{-T} \lambda_\mathcal{T}(-\bar{H}_\mathcal{T}^{-1}\bar{H}_\mathcal{C} q_\mathcal{C}[k] + \bar{H}_\mathcal{T}^{-1}\bar{d}[k]) = 0. \tag{8}$$

Now, due to mass conservation in the network, the flow out of the network must be equal to the flow into the network. Therefore there can only be $(n-1)$ independent nodal flows in the network, leading to the relation,

$$d_n[k] = -\sum_{i=1}^{n-1} d_i[k]. \tag{9}$$

The following can be stated about the incidence matrix, which is proven by Kallesøe et al. (2015).

Lemma 1 (Kallesøe et al., 2015)**.** *Let $\mathcal{T}$ be a directed tree with the incidence matrix $H_{\mathcal{T}}$ and the reduced incidence matrix $\bar{H}_{\mathcal{T}}$ (without loss of generality assuming that the last row of $H_{\mathcal{T}}$ has been removed to obtain $\bar{H}_{\mathcal{T}}$). The reduced incidence matrix is invertible since a tree is a connected graph with n-1 edges (Deo, 2017). Then the following holds*

$$H_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1} = \begin{bmatrix} I_{n-1} \\ -\mathbb{1}^T \end{bmatrix}, \tag{10}$$

where $\bar{H}_{\mathcal{T}}^{-1}$ denotes the right inverse of $\bar{H}_{\mathcal{T}}$.

Now, restricting to the spanning tree, $\mathcal{T}$, part of (5) gives,

$$H_{\mathcal{T}}^T p[k] = \Delta p_{\mathcal{T}}[k] = \lambda_{\mathcal{T}}(q_{\mathcal{T}}[k]) - H_{\mathcal{T}}^T z, \tag{11}$$

where $\Delta p_{\mathcal{T}}[k]$ is the vector of drop in pressure across the edges of $\mathcal{T}$. Multiplying (11) by $\bar{H}_{\mathcal{T}}^{-T}$ from the left gives,

$$\bar{H}_{\mathcal{T}}^{-T} H_{\mathcal{T}}^T \begin{bmatrix} \bar{p}[k] \\ p_n[k] \end{bmatrix} = \bar{H}_{\mathcal{T}}^{-T} \lambda_{\mathcal{T}}(q_{\mathcal{T}}[k]) - \bar{H}_{\mathcal{T}}^{-T} H_{\mathcal{T}}^T \begin{bmatrix} \bar{z} \\ z_n \end{bmatrix}, \tag{12}$$

where $p[k] = \begin{bmatrix} \bar{p}[k]^T & p_n[k] \end{bmatrix}^T$ (with $\bar{p}$ denoting non-reference pressures and p_n representing reference node pressures) and $z = \begin{bmatrix} \bar{z}^T & z_n \end{bmatrix}^T$. Applying Lemma 1 to (12), and furthermore using (4) gives an expression of $\bar{p}$ in terms of q_C.

$$\bar{p}[k] = \bar{H}_{\mathcal{T}}^{-T} \lambda_{\mathcal{T}}(-\bar{H}_{\mathcal{T}}^{-1} \bar{H}_C q_C[k] + \bar{H}_{\mathcal{T}}^{-1} \bar{d}[k]) - (\bar{z} - \mathbb{1} z_n) + \mathbb{1} p_n[k] \tag{13}$$

With that, the graph theory based hydraulic model of a water distribution network is given by the pressure equation (13) and the flow equation (8). Further, in Section 3.1, this model is reduced for a specific topology of water networks which will serve as the foundation for the leakage detection and localization algorithm.

3.1. Reduced order network model

In this section, the reduced order model, which has been previously derived by Rathore et al. (2022), under certain assumptions and considerations is presented.

The water networks considered in this project are District Metering Areas (DMAs) with multiple supply points but without an elevated reservoir. Let s be the number of supply nodes, then the independent nodal flows can be partitioned into a vector of supply nodal flows, $d_s[k] \in \mathbb{R}^s_{\geqslant 0}$, and a vector of non-supply nodal flows, $d_c[k] \in \mathbb{R}^{(n-s)}_{\leqslant 0}$, and this can be given as,

$$d[k] = \begin{bmatrix} d_c[k] \\ d_s[k] \end{bmatrix}. \tag{14}$$

As mentioned before, the derived model is intended to model the behaviour of a single DMA. The DMA is assumed to be sectionalized such that the consumers in the DMA are of the same type, i.e. either residential or industrial. Moreover, it is assumed that the same type of consumers behaves similarly, which is a standard assumption in water network analysis tools (Rossman et al., 2000). With this, the same consumer demand pattern can be assumed for all the consumers, though scaled for each consumer independently depending on size. Furthermore, we consider that one of the supply nodes is pressure controlled and all the other supply nodes are flow controlled.

Remark 1. A similar control structure is present at the water distribution network in Randers, Denmark operated by the company Verdo and is common in the operation of water networks.

Also, the pressure controlled supply node is denoted as the nth node. The flow controllers are designed to ensure that the distribution of flow from the supply nodes is always constant. These conditions are formally stated in Assumption 1.

Assumption 1. The distribution between the $n - s$ non-supply flows, $d_c[k]$, is fixed in time, i.e. $\exists\ w_c \in \mathbb{R}^{n-s}_{\geqslant 0}$, with the property $\sum_{i=1}^{n-s} w_i = 1$, such that,

$$d_c[k] = -w_c \gamma[k], \tag{15}$$

where $\gamma[k]$ is the total supply flow (or the total consumer demand). Also, the distribution between the $s - 1$ supply flows of $\bar{d}_s[k]$ is fixed in time, i.e. $\exists\ w_s \in \mathbb{R}^{s-1}_{\leqslant 0}$, with the property $\sum_{i=n-s+1}^{n-1} w_i = -1 + \kappa$, such that,

$$\bar{d}_s[k] = -w_s \gamma[k], \tag{16}$$

where, κ is the ratio of the reference node supply flow, $d_n[k]$, to the total supply flow, $\gamma[k]$, and is given as,

$$\kappa = \frac{d_n[k]}{\gamma[k]}. \tag{17}$$

With that, the non-reference nodal flows can be represented as,

$$\bar{d}[k] = -w\gamma[k] = -\begin{bmatrix} w_c \\ w_s \end{bmatrix} \gamma[k] \quad ; \quad \sum_{i=1}^{n-1} w_i = \kappa\ . \tag{18}$$

Local flow controllers are implemented at the non-reference supply nodes with the set-point $\bar{d}_s^*[k] = -w_s \gamma[k]$ to maintain a fixed distribution of the supply flows. As mentioned earlier, the leakage localization approach presented here is an extension of Rathore et al. (2021), which was designed for networks with a single supply node, whereas this work allows multiple supply nodes. The localization approach is based on a reduced order model which requires the distribution of the nodal flows to be constant. With the addition of multiple supply nodes, the flow from these nodes also needs to satisfy this requirement, and therefore a constraint on flow control of the supply nodes is imposed. Lemma 2 demonstrates the implication of this fixed distribution. Moving forward we set the pressure controlled supply node as the reference node which, as mentioned before, is also denoted as the nth node. With that (9) can be given as,

$$d_n[k] = -\left(\sum_{i=1}^{n-s} d_i[k] + \sum_{i=n-s+1}^{n-1} d_i[k] \right) = -\mathbb{1}^T \bar{d}[k]. \tag{19}$$

Now, the flow from the flow controlled supply points can be treated as negative nodal consumer flows. For a network with pressure control at the single reference node, and where the non-reference nodal flows (the consumer flows and the non-reference node supply flow) are a linear function of the total supply flow γ, i.e. Assumption 1, the Lemma 2 originally proven for water networks with single supply point (Kallesøe et al., 2015), holds for networks with multiple supply points.

Lemma 2 (Kallesøe et al., 2015)**.** *Under Assumption 1, for a given distribution of nodal flows at the non-reference nodes, w, such that $\bar{d}[k] = -w\gamma[k]$ there exists a unique vector a, such that $q[k] = a\gamma[k]$, where $q[k]$ is the vector of edge flows through the network.*

Lemma 2 implies that for a constant distribution of non-reference nodal flows, w, there exists a constant distribution of edge flows in the network, a. This distribution is independent of time instance, k, and the magnitude of the total demand $\gamma[k]$.

Now, with λ_i being a homogeneous function with degree 2, i.e. $\lambda_i(a_i x) = \lambda_i(a_i)x^2$ for $x \geqslant 0$, and Lemma 2, the network pressure model, (13), can be given as,

$$\bar{p}[k] = \gamma[k]^2 g(a_C, w) + b p_n[k] + c\ , \tag{20}$$

where, the vector a_C is part of a from Lemma 2 that relates to the chord flows of the spanning tree and the vector function and vectors g, b, and c are given as,

$$g(a_C, w) = \bar{H}_{\mathcal{T}}^{-T} \lambda_{\mathcal{T}}(-\bar{H}_{\mathcal{T}}^{-1} w - \bar{H}_{\mathcal{T}}^{-1} \bar{H}_C a_C)$$
$$b = \mathbb{1} \quad , \quad c = \mathbb{1} z_n - \bar{z}\ .$$

Note that the vectors w and a are time-invariant by Assumption 1 and Lemma 2 respectively.

Similarly, the model for flows in the network is given as,

$$h(a_C, w)\gamma[k]^2 = 0\,, \tag{21}$$

where,

$$h(a_C, w) = \lambda_C(a_C) - \bar{H}_C^T \bar{H}_{\mathcal{T}}^{-T} \lambda_{\mathcal{T}}(-\bar{H}_{\mathcal{T}}^{-1} w - \bar{H}_{\mathcal{T}}^{-1} \bar{H}_C a_C)\,.$$

(20) and (21) form the reduced order model for the water network DMA with multiple supply points under the explicit control structure considered in this work. Further, Section 4 presents the leakage diagnosis framework, which is based on the reduced order model mentioned earlier.

4. Leakage detection and localization

Leakage detection and localization presented in this work are based on pressure residuals. The pressure residuals are generated using a data-driven model and then compared to the sensitivity matrix for leakage localization. In this section, a model for pressure residual is presented which is used to obtain the sensitivity matrix and the expected residual signatures. Further, based on the same the leakage diagnosis framework is presented.

4.1. Pressure residual model

The pressure residual model is derived using first-order Taylor expansion where the change in pressure due to leakage is considered a small variation around the nominal pressure in the no-leakage scenario. Further in the paper *nominal* is used to refer to the no-leakage scenario. Similar models are presented by Rathore et al. (2021) and Rathore et al. (2022).

Let $\bar{D}[k]$ and $\bar{Q}[k]$ be the nodal flows at the non-reference nodes and the nominal flows in the edges respectively. Likewise, let W be the nominal distribution of the nodal flows such that, $\bar{D}[k] = W\Gamma[k]$ and A be the nominal distribution of edge flows such that, $Q[k] = A\Gamma[k]$, where $\Gamma[k]$ is the nominal total supply flow. Leakages in the network can be modelled as variations around the nominal nodal flows as,

$$\bar{d}[k] = \bar{D}[k] + \delta\bar{d}[k], \tag{22}$$

where $\bar{d}[k]$ is the actual non-reference nodal flows and $\delta\bar{d}[k] \in \mathbb{R}^{n-1}$ represents the leakages the network. Leakages would also result in variations in the actual distribution of the nodal flows and the distribution of edge flows, which is given by,

$$a[k] = A + \delta a[k] \quad , \quad w[k] = W + \delta w[k]\,, \tag{23}$$

where $\delta w[k]$ represents the variation in the distribution of the nodal flows and $\delta a[k]$ represents the variation in the distribution of edge flows from their respective nominal values due to leakage $\delta\bar{d}[k]$. With these small variations around the nominal values, the pressure variation, $\delta p[k]$, around the nominal pressure, $\bar{P}[k]$, due to leakage is given by Taylor expansion of (20) in (24).

$$\begin{aligned}\bar{P}[k] + \delta\bar{p}[k] \;=\; & \gamma[k]^2 g(A_C, W) + bp_n[k] + c \\ & + \gamma[k]^2(\partial_{a_C} g|_{A_C,W}\delta a_C[k] + \partial_w g|_{A_C,W}\delta w[k]) + \mathcal{O},\end{aligned} \tag{24}$$

where the $\mathcal{O}$ represents in the higher order terms.

Further, ignoring the higher-order terms and considering the nominal pressure,

$$\bar{P}[k] = \gamma[k]^2 g(A_C, W) + bp_n[k] + c, \tag{25}$$

the pressure variation can be approximated by,

$$\delta\bar{p}[k] \approx \gamma[k]^2(\partial_{a_C} g|_{A_C,W}\delta a_C[k] + \partial_w g|_{A_C,W}\delta w[k]). \tag{26}$$

Note that this pressure variation is nothing but an approximate model of pressure residuals.

Similarly, from a first-order Taylor expansion of (21), we obtain an expression for the variation of the edge flows distribution δa_C.

$$\gamma[k]^2(h(A_C, W) + \partial_{a_C} h|_{A_C,W}\delta a_C[k] + \partial_w h|_{A_C,W}\delta w[k]) \approx 0$$

As presented by Rathore et al. (2022) the expression for $\delta\bar{p}$ is given as,

$$\delta\bar{p}[k] = -\gamma[k]^2 S\delta w[k]\,, \tag{27}$$

where S is termed as the resistance matrix and is given by,

$$S = (\bar{H}_{\mathcal{T}}\partial\lambda_{\mathcal{T}}^{-1}\bar{H}_{\mathcal{T}}^T + \bar{H}_C\partial\lambda_C^{-1}\bar{H}_C^T)^{-1}. \tag{28}$$

Note that the S matrix only depends on the nominal distribution of the non-reference nodal flows, W.

Remark 2. The part of W corresponding to non-supply nodes can be estimated from billing data and the part corresponding to the supply nodes is already known and fixed by the virtue of the supply flow control.

With the structure of $\lambda_i(\cdot)$, (6), $\partial\lambda_{\mathcal{T}}$ and $\partial\lambda_C$ are diagonal matrices implying that S is symmetric, and $\partial\lambda_i > 0$ for $q_i \neq 0$. In a real-life water network, $\partial\lambda_C$ and $\partial\lambda_{\mathcal{T}}$ are always non-singular, as all edges are pipes with an inherent flow resistance.

With mass conservation in the network, (19),

$$D_n[k] = -\mathbb{1}^T\bar{D}[k] \quad , \quad d_n[k] = -\mathbb{1}^T\bar{d}[k]. \tag{29}$$

Further, from Assumption 1 with flow control at all the non-reference supply nodes, the flow at the reference supply node can be represented as a fixed ratio, κ, of the total supply flow, $\gamma[k]$ (and $\Gamma[k]$), as,

$$D_n[k] = \kappa\Gamma[k] \quad , \quad d_n[k] = \kappa\gamma[k]. \tag{30}$$

Therefore,

$$\kappa\Gamma[k] = -\mathbb{1}^T\bar{D}[k] \quad , \quad \kappa\gamma[k] = -\mathbb{1}^T\bar{d}[k]. \tag{31}$$

With that rewriting (22) gives,

$$\Gamma[k] - \gamma[k] = \frac{1}{\kappa}\mathbb{1}^T\delta\bar{d}[k]\,. \tag{32}$$

Now, the nominal non-reference nodal flow $\bar{D}$ and the actual non-reference nodal flow $\bar{d}$ with a leakage $\delta\bar{d}$ is described by,

$$\bar{D}[k] = -W\Gamma[k] \quad , \quad \bar{d}[k] = -(W + \delta w[k])\gamma[k]\,. \tag{33}$$

Substituting (33) in (22) and solving for δw using (32) leads to the following relation between $\delta w[k]$ and $\delta d[k]$,

$$\delta w[k] = -\frac{1}{\gamma[k]}\left(I - \frac{1}{\kappa}W\mathbb{1}^T\right)\delta\bar{d}[k]. \tag{34}$$

The following can be stated about the matrix $(I - \frac{1}{\kappa}W\mathbb{1}^T)$, which has been proven by Rathore et al. (2022).

Lemma 3 (*Rathore et al., 2022*). *Let $M = I - \frac{1}{\kappa}W\mathbb{1}^T$, where $\sum_i W_i = \kappa$, then M has a non-trivial kernel,*

$$ker(M) = span\{W\}\,.$$

Now, substituting (34) in (27), gives the relation between leakage $\delta\bar{d}[k]$ and the pressure variation $\delta\bar{p}[k]$ in (35).

$$\delta\bar{p}[k] = \gamma[k]S\left(I - \frac{1}{\kappa}W\mathbb{1}^T\right)\delta\bar{d}[k]. \tag{35}$$

Assuming, in a real-life scenario there is always a flow in all the pipes, i.e. $\forall i,\ q_i \neq 0$, making S full rank. Therefore with Lemma 3, $\delta\bar{p}[k]$ is zero only when $\delta\bar{d}[k] = 0$ or $\delta\bar{d}[k] \in \text{span}\{W\}$. Here, $\delta\bar{d}[k]$ is used to model leakages in the network, hence $\delta\bar{d}[k] = 0$ implies no leakage in the network. Further, Lemma 3 implies that as long as the non-reference nodal flow distribution between the nodes, in case of a set of leaks, is different from the nominal distribution W, the pressure residuals $\delta\bar{p}[k]$ would be non-zero and hence the leakages would be

visible in the pressure residuals. Therefore, in practice, a leakage at a single node would always be detectable. In the following sections, leakage detection, localization and identification based on the reduced order model and the pressure residual model are presented, assuming that a leakage appears only one node at a time.

4.2. Leakage detection

The leakage diagnosis framework presented in this work is based on pressure residuals. Therefore we start with the generation of residuals, on which a decision system is implemented to detect if leakage has occurred or not.

Generally, network pressure is only measured at a few of the nodes, which is a subset of $\bar{p}$. The pressure residuals are also only generated for these nodes, by comparing the measured pressure to the estimated pressure. The pressure and flow at the supply nodes are also typically measured in a water network. Considering that, given the pressure at the reference node, $p_n[k]$, and the total supply flow, $\gamma[k]$, the non-reference nodal pressures in the network under nominal conditions can be estimated using (25). The estimated nominal nodal pressure at the ith node can be given as,

$$\hat{P}_i[k] = \alpha_i \gamma[k]^2 + \beta_i + p_n[k], \tag{36}$$

where, from (20), α_i is the ith element of $\bar{H}_{\mathcal{T}}^{-T}\lambda_{\mathcal{T}}(\bar{H}_{\mathcal{T}}^{-1}W - \bar{H}_{\mathcal{T}}^{-1}\bar{H}_C A_C)$ and β_i is the ith element $-(\bar{z} - \mathbb{1}z_n)$.

With the estimated nominal distribution of the non-reference nodal flows, W, the parameters α_i and β_i can be estimated with the help of standard hydraulic network models such as EPANET (Rossman et al., 2000). However, typically the pressure residuals due to leakages are small, which would lead to high precision requirements for the standard hydraulic models. In our approach, we relax these requirements by using a data-driven model for the residual generation. Assuming nominal conditions, given time series data of measured network pressure at the ith node, $P_i[k]$, $\gamma[k]$ and $p_n[k]$, the α_i and β_i parameters can be identified using linear regression (Madsen, 2007).

Once the parameters are identified over a window of data, the nominal pressure at the measured nodes, after the window, can be estimated using the measured $p_n[k]$ and $\gamma[k]$ with (36). Further, pressure residuals can be calculated as,

$$r_i[k] = p_i[k] - \hat{P}_i[k], \tag{37}$$

where p_i is the measured pressure at the ith node and $\hat{P}_i$ is its estimated value in the non-leaking case. The model parameters can be updated continuously over a moving window, which would ensure that changes in the system are incorporated into the model. However, a burst leakage in the network would be a sudden change and would be detected in the residuals.

Under ideal conditions, with the known fixed distribution of consumer demands and no noise, the zero value of r would indicate no leakage and a non-zero value would indicate that leakage has occurred. However, in practical conditions, the residual signal is impacted by noise and therefore a statistical change detector needs to be employed to detect leakage. Here, we propose to use the Generalized Likelihood Ratio (GLR) test (Blanke, Kinnaert, Lunze, Staroswiecki, & Schröder, 2006) for detecting a change in the mean of the residuals. The mean value of the residual would be zero during the normal operation of the network and when a leakage has occurred it would deviate from zero. The GLR decision function, $\phi_i[k]$, for the ith residual is given by,

$$\phi_i[k] = \frac{1}{2\sigma_i^2} \max_{k-M_w+1 \leqslant j \leqslant k} \frac{1}{k-j+1}\left(\sum_{s=j}^{k} r_i(s)\right)^2, \tag{38}$$

where σ_i is the standard deviation of the residual signal, M_w is the sliding window and r_i is the ith residual signal. When any one of the GLR decision functions crosses the threshold value, thr; i.e. if $\phi_i(k) > thr_i$ for any i, a leakage detection alarm is generated. In the following section, we use the residual vector r for leakage localization.

4.3. Leakage localization

A leakage can be modelled as an unexpected nodal flow. We assume leakages only occur at non-supply nodes and, as mentioned before, it only occurs at one node at a time. Therefore, a leakage at node $l \in \mathbb{Z} : l \in [1, (n-p)]$ can be represented as

$$\delta\bar{d}[k] = \begin{bmatrix} e_l \\ w_s \end{bmatrix}\zeta[k] = -\begin{bmatrix} e_l \\ w_s \end{bmatrix}|\zeta[k]|, \tag{39}$$

where e_l represents the location of the leakage with $e_l \in \{0,1\}^{(n-p)}$ being a unit vector with 1 at lth position and w_s is the distribution between the $s-1$ non-reference supply flows. And $\zeta[k] < 0$ is the magnitude of the leakage. With this model of leakage, the pressure variation (or residual) model (35) can be given as,

$$\delta\bar{p}[k] = -\gamma[k]S\left(I - \frac{1}{\kappa}W\mathbb{1}^T\right)\begin{bmatrix} e_l \\ w_s \end{bmatrix}|\zeta[k]|. \tag{40}$$

Further, we compare the measured residual, (37), to the pressure residual model to localize a leakage. (40) includes pressure residuals for all the nodes, and from that to extract nodes at which the pressure is measured we use a binary matrix F. With that, the pressure residual model or the expected residual for leakage at lth node is given as,

$$\hat{r}_l[k] = G\begin{bmatrix} e_l \\ w_s \end{bmatrix}\gamma[k]|\zeta[k]|. \tag{41}$$

Here, G is named the sensitivity matrix and is given by,

$$G = -FS\left(I - \frac{1}{\kappa}W\mathbb{1}^T\right), \tag{42}$$

G only depends on the distribution of the nominal non-reference nodal flows, W, which as mentioned before can be estimated from the billing data.

From (41) it can be seen, $G\begin{bmatrix} e_l \\ w_s \end{bmatrix}$ gives the direction of the residuals, $r_l[k]$. Moreover, the direction is neither impacted by total supply flow, $\gamma[k]$, nor by the magnitude of the leakage $\zeta[k]$. To localize leakage, we compare the direction of the actual residual, $r[k]$ from (37), with the set of expected residual models described by $\hat{r}_i$ $i = 1, \ldots, n-p$ from (41). The comparison is made using the angle between the vector, leading to the following decision signal,

$$\psi_i(r[k]) = \frac{\left\langle r[k], G\begin{bmatrix} e_i \\ w_s \end{bmatrix}\right\rangle}{|r[k]|\left|G\begin{bmatrix} e_i \\ w_s \end{bmatrix}\right|}. \tag{43}$$

This approach for comparing measured residual to leakage signature has been previously presented by Perez et al. (2014) and Rathore et al. (2022).

(43) implies $-1 \leqslant \psi_i \leqslant 1$ and the leakage node can be estimated by,

$$\hat{l} = \arg\max_i \{\psi_i\}_{i=1,\ldots,n-1}. \tag{44}$$

The leakage indicator can also be obtained by truncating the decision signal $\psi_i(r)$ to be zero or larger than zero as,

$$\mu_i[k] = \max\left\{\psi_i(r[k]), 0\right\}_{i=1,\ldots,n-1}, \tag{45}$$

where $\mu_i \in [0,1]$ is the leakage indicator for the ith node.

In a real-life water network, with a few numbers of pressure sensors compared to the number of nodes, multiple nodes would have the leakage indicator value close to $\hat{l}$'s value (the maximum value), $\mu_{\hat{l}}[k]$. Therefore, instead of indicating a leakage at one node, the leakage indicator would point towards a set of nodes, having a higher likelihood of leakage with a μ value close to 1. With that, the leakage indicator signal leads the utility to an area of the network that should be inspected for leakage. Further, in Section 5 the contamination mitigation control is developed using the information of the identified leaking node.

5. Contamination mitigation control for water distribution network under leakage

Using the terminology presented by Sadiq et al. (2006), the entry of contaminants into a water distribution system relies on three factors: the presence of a pathway, a driving force, and a source of contamination. In our case the pathway would be the leakage point, the driving force would be the low/negative pressure in the network and the source of contamination would be the environment around the leakage point. When there is a leak in a water network, the water flows out of the network into the surrounding. This could result in a pool of water forming around the leakage point, which may become contaminated by various pathogens. Further, when the pressure drops inside the network, this external fluid could lead to the intrusion of pathogens potentially contaminating the network's water (Mora-Rodríguez et al., 2015). This negative pressure or a pressure drop can arise from pressure transients within the pipeline, which results from sudden variations in water velocity caused by events such as leakage, abrupt changes in demand, or unregulated pump starting and stopping. These pressure transients have the potential to induce negative pressures even in a pressurized distribution system (Karim et al., 2003). The risk of contamination increases significantly when a garbage disposal area or wastewater line is in close proximity to the leakage point (ASCE, 2004; Deng, Jiang, & Sadiq, 2011). Mora-Rodríguez et al. (2015) models contaminant scenarios for different types of pipe failures with different shapes and sizes of leakage points in buried and unburied pipes.

To prevent this contamination we propose implementing a control strategy that directs the flow of water towards the leak area and isolates the affected nodes from the rest of the network. With this approach, the network continuously flushes water through the leakage point and the flows in the network cannot carry any potential contaminants around the network. The idea to flush water out of the leakage point may initially seem counterproductive, as it could potentially exacerbate the issue or cause further damage to the network. However, the objective here is not to flush any contaminant out of the network but to prevent entry and spread of contaminant from the leakage point. Therefore, in this work, the contamination mitigation control is formulated as an optimization problem with the objective to maintain the nominal operational pressure in the network and the flow directions are incorported as constraints of the optimization problem. This ensures that the flow is towards the leakage area to prevent the spread of the contaminant in the network. As a result, the likelihood of contamination spread is significantly reduced, while the cost of water wastage or damage to the network is kept to a minimum. Once the leakage alarm is generated and the leakage node set is identified, the contamination mitigation control is activated.

In this section, we present an algorithm for identifying the edges that need to be isolated in order to cut off leaking nodes from the rest of the water network. We also provide a control strategy for minimizing the risk of contamination from the leak. The first step in the process is to use the algorithm to determine the edges and flow direction that will allow us to isolate the leaking nodes. Based on this information, we propose an optimization problem to mitigate contamination from the leak.

5.1. Leakage nodes isolation

To isolate a leaking node from the rest of the network, we need to identify the edges that connect the leak to the network. In our case, where the leakage localization algorithm identifies a set of nodes as the source of the leak, viz. leakage node set, we would isolate the entire set from the network using the edges connecting that set to the network. Again, these would be the edges where the flow direction would be towards the leakage node set. To provide a clearer understanding of this process, we will use the graph shown in Fig. 1 as an example.

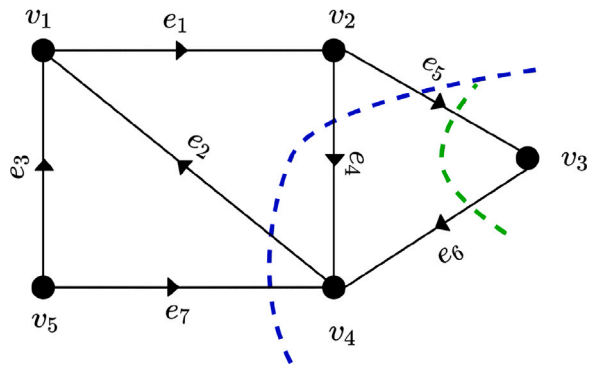

Fig. 1. Graphical representation of a small water network used as an example to illustrate the concept of isolating leakage nodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Let Fig. 1 be a graphical representation of a small water network with 5 nodes, $\{v_1, \dots, v_5\}$, and 7 edges, $\{e_1, \dots, e_7\}$. In this graph, if node v_3 is identified as the leaking node, then to isolate it from the rest of the network, the flow direction in the edges e_5 and e_6 needs to be towards v_3. Specifically, the flow direction in e_5 should be from v_2 to v_3 and in e_6 it should be from v_4 to v_3. If, on the other hand, both nodes v_3 and v_4 are identified as the leaking node set, then the flow direction in the edges e_2, e_4, e_5 and e_7 needs to be towards v_3 and v_4. It is worth noting that there is no constraint on the flow direction in the edge e_6, which connects the two leaking nodes.

Now, to identify these edges we use cuts in the graph. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a directed graph with $\mathcal{V}$ being the set of nodes and $\mathcal{E}$ being the set of edges of the graph. Further, let $\mathcal{V}_o$ be a non-empty subset of $\mathcal{V}$, then the set of edges connecting $\mathcal{V}_o$ to $\mathcal{V} - \mathcal{V}_o$ is a cut. For a single node v_o, the cut from v_o to $\mathcal{V} - \{v_o\}$ can be given by the row of the incidence matrix H corresponding to the node v_o (Thulasiraman & Swamy, 2011). Using that, in a water network, the cuts or the set of edges to isolate a node v_o from the rest of the network is given as,

$$\mathcal{E}_{v_o} = diag(H_{v_o})^{|\cdot|} e, \tag{46}$$

where H_{v_o} is the row of the incidence matrix H corresponding to the node v_o and e is the vector of edges of the graph.

To illustrate this concept, we again use the graph shown in Fig. 1. The incidence matrix for this graph can be calculated using (1), and it is given as follows,

$$H = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array} \begin{array}{c} \begin{array}{ccccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 & e_7 \end{array} \\ \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}. \tag{47}$$

Now, as mentioned before, to isolate v_3 the cuts or the edges with flow direction towards v_3 are e_5 and e_6. In Fig. 1, these edges are represented by a green dot curve over them. This is also seen in the 3rd row of the incidence matrix with non-zeros entries in the 5th and 6th column, which gives,

$$\mathcal{E}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{|\cdot|} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix} = \begin{bmatrix} e_5 \\ e_6 \end{bmatrix}. \tag{48}$$

This can be extended to the isolation of a set of nodes $\mathcal{V}_o$ as,

$$\mathcal{E}_{\mathcal{V}_o} = E^{|\cdot|} e, \tag{49}$$

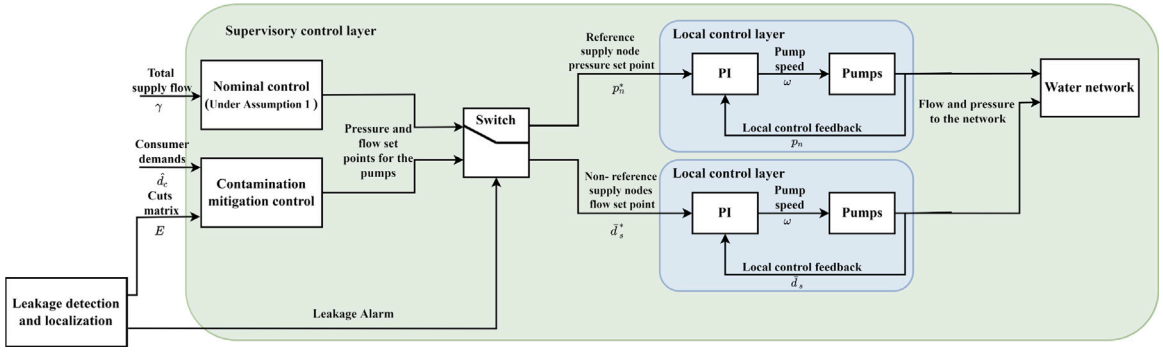

Fig. 2. The overall control structure proposed in this work.

where,

$$E = diag\left(\sum_{v_o \in \mathcal{V}_o} H_{v_o}\right). \tag{50}$$

Here we sum the row of the incidence matrix H corresponding to all the nodes $v_o \in \mathcal{V}_o$. The sum of the rows in the expression takes into account all the edges connecting the leaking nodes to the rest of the network. For the ith edge that connects two leaking nodes, the corresponding entries in the ith column of H will be 1 and -1 for the two nodes, respectively. These entries will cancel out when they are summed, ensuring that the edges connecting the leaking nodes to each other are not included in (49). This is necessary to ensure that only the edges that connect the leaking nodes to the rest of the network are considered.

Again, in Fig. 1, to isolate nodes v_3 and v_4 the cuts are represented by a blue dotted curve over them. The cut edges can be obtained by summing the 3rd and the 4th row of the incidence matrix, H and this is given as,

$$\mathcal{E}_{\{3,4\}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}^{|\cdot|} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix} = \begin{bmatrix} e_2 \\ e_4 \\ e_5 \\ e_7 \end{bmatrix}. \tag{51}$$

The cut orientation in E is from $\mathcal{V}_o$ to $\mathcal{V} - \mathcal{V}_o$ however, we want the flow direction in the cut to be from the network to the leakage node set, i.e. from $\mathcal{V} - \mathcal{V}_o$ to $\mathcal{V}_o$. Therefore, the constraint on flow direction in the cut is given as,

$$Eq[k] \leqslant 0. \tag{52}$$

This constraint on flow directions will ensure that any potential contaminant is not carried around the network. As previously stated, considering a potential leakage at nodes v_3 and v_4, the flow direction in the edges e_2, e_4, e_5 and e_7 needs to be towards v_3 and v_4. Therefore with respect to the direction of the edges in the graph, q_2 should be negative and q_4, q_5 and q_7 should be positive in order to prevent contamination. Further, in the following section, (52) is set as a constraint in the optimization problem.

5.2. Contamination mitigation control design

In this section, we present the contamination mitigation control design to ensure that the direction of flows in the cut edges for the leakage node set is towards the leakage node set. The control also needs to ensure that the consumer demands are met with the required pressure.

Therefore, we propose the control structure solution presented in Fig. 2 for the entire framework. The control is implemented in two layers, viz. the supervisory control layer and the local control layer. The supervisory control layer consists of the nominal control and the contamination mitigation control. As previously mentioned, the reference supply node is pressure controlled and the rest are flow controlled. Therefore, the output of both the controls are pressure set-points for the reference supply node and flow set-points for the rest of the supply node. Outputs from both controls go into a switch, under nominal conditions the output of the switch is the set-point from the nominal control and when a leakage alarm is generated the output switches to the set-point from the contamination mitigation control. The input to the nominal control is the total supply flow and as per Assumption 1 based on the fixed distribution of supply flows it generates the set-points. The input to the contamination mitigation control is the predicted consumer demands and the E matrix from (52). The contamination mitigation control solves an optimization problem to generate the set-points. The supply here is assumed to be from pumping stations, and the set-points are sent to the local control in the pumping stations. The local control could be PI control or a state space-based control. The design of local control is not part of this work and is simply assumed to be present. The local control regulates the angular velocity ω of the pumps to achieve desired flow and pressure set-points.

The contamination mitigation control solves an optimization problem with a cost function to be minimized under some constraints. Here, the objective is to maintain the network pressure at the desired level while fulfilling the constraint (52). The network pressure can be characterized as the pressure at all nodes in the network or a subset of nodes, which may be sufficient to provide an accurate estimation of pressure in a particular area. Selecting a subset of the nodes could also help in reducing the computational load of the optimization problem. The operator or utility will have the discretion in selecting this subset of nodes. Henceforth the nodes part of this subset are called critical nodes. With that, the cost function is formulated as,

$$\mathcal{J}[k] = (p_c[k] - p_c^*[k])^T (p_c[k] - p_c^*[k]), \tag{53}$$

where p_c is the network pressure at critical nodes and p_c^* is the desired pressure set-point.

The critical node pressures can also be represented in terms of non-reference node pressures as,

$$p_c[k] = F_c \bar{p}, \tag{54}$$

where F_c is a binary matrix to extract the critical node pressures. In the case of pressure control at all the nodes, the matrix F_c would be an identity matrix. With that, Eq. (53) is given as,

$$\mathcal{J}[k] = (F_c \bar{p}[k] - p_c^*[k])^T (F_c \bar{p}[k] - p_c^*[k]). \tag{55}$$

Further, we present the constraints of the optimization problem.

5.2.1. Constraints

We start with the flow direction constraint in the cut edges, where the flow direction in the cut edges should be towards the leakage node set, and is given by (52). However, under some conditions, it might not be possible to fulfil this constraint and that would lead to an infeasible optimization problem. To avoid that we soften this constraint by adding a slack variable, ϵ, and the softened constraint is given as,

$$Eq[k] \leqslant \epsilon[k], \tag{56}$$

ϵ is defined such that it is non-zero only if (52) is violated and when it is non-zero it is heavily penalized in the cost function.

Further, the capacity of the pumps also forms a practical constraint as they can only deliver a maximum pressure p_s^{max} and flow d_s^{max}. This constraint in the optimization problem is formulated as,

$$0 \leqslant F_s \bar{p}[k] \leqslant p_s^{max}, \tag{57}$$

and,

$$0 \leqslant d_s[k] \leqslant d_s^{max}, \tag{58}$$

where F_s is a binary matrix to extract the supply node pressures.

Lastly, there is an upper and lower bound on the pressure at the critical nodes, which is formulated as,

$$p_c^{min} - \xi \leqslant F_c \bar{p}[k] \leqslant p_c^{max} + \xi, \tag{59}$$

where p_c^{min} is the minimum acceptable pressure and p_c^{max} is the maximum acceptable at the critical node. Again, ξ is the slack variable for (59) to avoid the infeasibility of the optimization problem. The actuators of the network are the pumps or supply points, and since control is only over the supply pressure and flow, there may be cases where it is not possible to direct all the flows towards the leakage area while satisfying the other constraints. Under these conditions, the optimization problem becomes infeasible. To avoid such infeasibility, slack variables ϵ and ξ are added to the constraints. Slack variables allow the constraints to be violated by a minimum amount in cases of infeasibility. In networks where pumping stations are located nearby or in networks with a tree-like structure, this control strategy may not be effective as it may not be possible to change and direct the flow towards the leakage area with pump control alone. However, even in these cases, the contamination mitigation control can still provide the best possible solution to limit the spread of the contaminant with the help of slack variables. Note that slack variables are not added to the pump capacity constraints, as those are physical constraints that cannot be violated.

5.2.2. The constrained optimization problem

Now, with the constraints and cost function, the optimization problem can be formulated as,

$$\min_{\bar{d}_s[k], p_n[k], \epsilon[k], \xi[k]} (F_c \bar{p}[k] - p_c^*[k])^T (F_c \bar{p}[k] - p_c^*[k]) + \epsilon[k]^T \mathcal{R} \epsilon[k] + \xi[k]^T \mathcal{Q} \xi[k] \tag{60}$$

subject to

$$\lambda_C(q_C[k]) - \bar{H}_C^T \bar{H}_\mathcal{T}^{-T} \lambda_\mathcal{T}(-\bar{H}_\mathcal{T}^{-1} \bar{H}_C q_C[k] + \bar{H}_T^{-1} \begin{bmatrix} d_c[k] \\ \bar{d}_s[k] \end{bmatrix}) = 0, \tag{61a}$$

$$\bar{p}[k] = \bar{H}_T^{-T} \lambda_\mathcal{T}(-\bar{H}_\mathcal{T}^{-1} \bar{H}_C q_C[k] + \bar{H}_T^{-1} \begin{bmatrix} d_c[k] \\ \bar{d}_s[k] \end{bmatrix}) - (\bar{z} - \mathbb{1} z_n) + \mathbb{1} p_n[k], \tag{61b}$$

$$Eq[k] \leqslant \epsilon[k], \tag{62a}$$

$$0 \leqslant F_s \bar{p}[k] \leqslant p_s^{max}, \tag{62b}$$

$$0 \leqslant d_s[k] \leqslant d_s^{max}, \tag{62c}$$

$$p_c^{min} - \xi[k] \leqslant F_c \bar{p}[k] \leqslant p_c^{max} + \xi[k]. \tag{62d}$$

The cost function, (60), of the optimization problem comes from (53) and added to it is the cost of the slack variables. $\mathcal{R}$ and $\mathcal{Q}$ are the weight on the cost associated with the slack variables. The values of $\mathcal{R}$ and $\mathcal{Q}$ are to be kept high to ensure that whenever $\epsilon[k] \neq 0$ or $\xi[k] \neq 0$, they are heavily penalized. The decision variables are the flow from the non-reference supply nodes, $\bar{d}_s$, the pressure at the reference node, p_n, and the slack variables, ϵ and $\xi[k]$. Further, the optimization problem is subjected to system model constraints given by (61). Finally, the constraints presented in Section 5.2.1 are included in (62). As stated before, the output, $\bar{d}_s[k]$ and $p_n[k]$, are set as the set-points for the local control.

Remark 3. In future work, the efficacy of contamination mitigation control under various leakage locations and magnitudes could be evaluated using the value of ϵ. It could also serve as a gauge to estimate the extent of the contamination in cases where complete elimination of contamination is unattainable.

5.3. Leakage diagnosis with contamination mitigation control framework

The complete leakage diagnosis with contamination mitigation control framework is presented in Algorithm 1. The input to the algorithm is the expected consumer demands and the parameters of the contamination mitigation control optimization problem are to be defined. Under nominal conditions, when the leakage alarm is inactive, the water network is controlled by the nominal control. The reference supply node pressure set-point is set as the nominal pressure set-point defined by the operator and the non-reference supply nodes flow set-points are set as per Assumption 1. With that, the leakage detection and localization algorithm is performed, which is presented in Algorithm 2. The input data for leakage detection and localization algorithm are the measured network pressures, reference supply node pressure and the total supply flow. Based on the measurements, the parameters of the network are identified and the pressure residuals are generated. Then a GLR decision function for each pressure residual is obtained and if any of the GLR decision functions crosses the predefined threshold, a leakage alarm is generated. The residuals are compared to the sensitivity matrix to calculate the leakage indicator for each node. And based on the leakage indicator, the leakage node set and the cut edges are identified. Further, when the leakage alarm is generated the control switches from nominal to contamination mitigation control. The contamination mitigation control optimization problem is solved with the expected consumer demands and the cut edges, identified by the leakage localization algorithm, to obtain the supply pressure and flow set-points. The reference supply node pressure set-point and the non-reference supply nodes flow set-points are set to the set-points obtained from the contamination mitigation control.

Algorithm 1 Leakage diagnosis with contamination mitigation control

Define Binary matrix to extract the critical node, F_c. Parameters for the optimization problem, p_c^*, p_s^{max}, d_s^{max}, p_c^{min}, p_c^{max}. Nominal control reference node pressure set-point, $\{p_n^*\}^{nom}$. Fixed supply flow distribution, w_s.
Input Expected consumer demands, $d_c[k]$.
while Leakage Alarm=0 **do**
▷ Nominal control
  Set reference supply node pressure set-point, $p_n^*[k] = \{p_n^*\}^{nom}$
  Set non-reference supply nodes flow set-point, $\bar{d}_s^*[k] = -w_s\gamma[k]$
  Apply supply pressure and flow set-points to the water network and perform Algorithm 2
end while
while Leakage Alarm=1 **do**
▷ Contamination mitigation control
  Solve optimization problem in Section 5.2.2 to obtain reference supply node pressure set-point $\{p_n^*\}^{con}[k]$ and non-reference supply nodes flow set-point,$\{\bar{d}_s^*\}^{con}[k]$ from the contamination mitigation control
  Set reference supply node pressure set-point, $p_n^*[k] = \{p_n^*\}^{con}$
  Set non-reference supply nodes flow set-point, $\bar{d}_s^*[k] = \{\bar{d}_s^*\}^{con}[k]$
  Output $p_n^*[k]$ and $\bar{d}_s^*[k]$
end while

Algorithm 2 Leakage detection and localization

Define Time window, M_w. GLR threshold, thr_i. Leak likelihood threshold, lk.
Input Set of pressure measured nodes, $\mathcal{N}_p$. Measured pressure in the network, $p_i[j]$, reference node pressure, $p_n[j]$ and total supply flow, $\gamma[j]$ over the period $j = (k - M_w) \cdots k$. Sensitivity matrix, G. Fixed supply flow distribution, w_s.
Initialize Leakage Alarm=0
while Leakage Alarm=0 **do**
  Obtain α_i and β_i, $\{\forall i | i \in \mathcal{N}_p\}$ for (36) with data $p_i[j]$, $p_n[j]$ and $\gamma[j]$ over the period $j = (k - M_w) \cdots (k - 1)$ using least square approximation
  Obtain $\hat{P}_i[k]$, $\{\forall i | i \in \mathcal{N}_p\}$ with α_i, β_i, $p_n[k]$ and $\gamma[k]$ using (36)
  Obtain $r_i[k]$, $\{\forall i | i \in \mathcal{N}_p\}$ with $\hat{P}_i[k]$ and $p_i[k]$ using (37)
  Obtain GLR decision function $\phi_i[k]$, $\{\forall i | i \in \mathcal{N}_p\}$ using (38)
  if $\phi_i[k] > thr_i$, for any $i \in \mathcal{N}_p$ **then**
    Set Leakage Alarm=1
    Obtain μ_i, $\forall i = 1 \cdots (n-1)$ using (43)–(45).
    Identify leakage node set $\mathcal{V}_o = \{v_i | \mu_i > lk\}$
    Obtain cut edge matrix, E, using (50)
  end if
end while
Output Leakage Alarm and E

The leakage diagnosis algorithm and the contamination mitigation control are implemented and tested as one framework on a laboratory water network setup. The implementation and the results are presented in Section 6.

6. Test results

In this section, we present the results of the performance evaluation of the leakage diagnosis with the contamination mitigation control framework. The evaluation was conducted in two parts. Firstly, a test was carried out at the Smart Water Infrastructures Laboratory (SWIL) located at Aalborg University, Denmark. Later, the framework was tested on a large-scale simulated benchmark network called L-town.

Table 1
Elevation of the nodes.

Node	v_1	v_2	v_3	v_4	v_5
Elevation [m]	0	0	0.9	0	0
Node	v_6	v_7	v_8	v_9	v_{10}
Elevation level [m]	0	0	0.9	0.9	0

6.1. Laboratory setup

The Smart Water Infrastructures Laboratory (SWIL) is a reconfigurable laboratory test bed composed of various modular components to emulate different parts of a water distribution network. These modules, such as the pumping station module, the pipe module, and the consumer station module, are equipped with some actuators and sensors for the control and monitoring of the network. A 3D illustration of one such module (pumping station) and a picture of the laboratory is shown in Fig. 3. The pumping station module from the 3D illustration can also be seen on the left side of the laboratory picture. The modules can be connected together using exterior valves and external pipes to create a customized water network. The networks built in laboratory setup are scaled, such that pressure and flow values are smaller compared to real-life networks and the time is accelerated. The Central Control Unit (CCU), which is a desktop computer, can remotely control the modules through Modbus TCP/IP communication. The CCU allows the implementation of various control and analysis algorithms using the MATLAB/Simulink program. Apart from the water distribution network, waste water systems and district heating networks can be emulated in the SWIL. More information about the SWIL and its applications can be found in Ledesma, Wisniewski, and Kallesøe (2021).

The laboratory water network setup and its corresponding network graph considered in this work are presented in Fig. 4. The network consists of 10 nodes, labelled as $v_1 \cdots v_{10}$, and 11 edges, labelled as $e_1 \cdots e_{11}$. The elevations of the nodes are listed in Table 1. The edges are the pipes of the network, and the length and diameter of each pipe are indicated in blue along the edges using the notation $x|y$, where x represents the length of the pipe in metres and y represents the diameter of the pipe in millimetres.

The network has two supply nodes, v_1 and v_{10}, which are connected to Pump 1 and Pump 10, respectively. Node v_{10} is set as the reference node and the pressure p_n is measured at this node. Subsequently, Pump 10 is pressure controlled and Pump 1 is flow controlled. The flows at both the supply nodes are measured and combined to give the total supply flow, γ. The network also has three consumers, connected at node v_3, v_8 and v_9. The consumers are labelled Consumer 3, 8 and 9, for the ease of associating with the nodes to which they are connected. In addition, three pressure sensors are used to measure the pressure in the network at nodes v_3, v_5, and v_7. The leakage detection and localization are carried out based on these pressure measurements.

6.2. Laboratory results

The consumer demands are emulated by controlling the flow control valves in the consumer module using individual PI controllers. In line with Assumption 1, the consumer demand pattern for all the 3 consumers is the same though scaled differently for each consumer. Here, the set-point for the PI controller is a periodic signal with a time period of 120 min. The set-points are generated by,

$$d_{cur}(t_h[k]) = a_0 + a_1 \cos \omega t_h[k] + b_1 \sin \omega t_h[k] + a_2 \cos \omega t_h[k] + b_2 \sin \omega t_h[k], \quad (63)$$

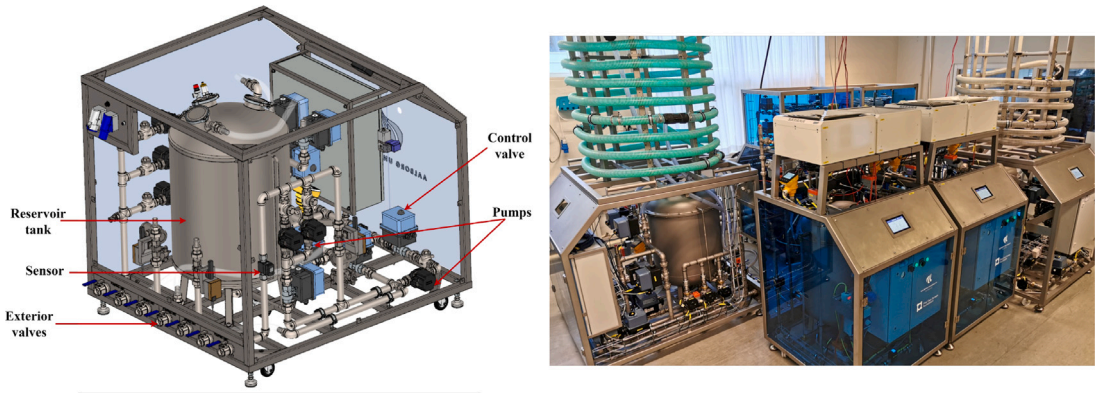

Fig. 3. A 3D illustration of the pumping station module from the SWIL and a picture of the laboratory setup.

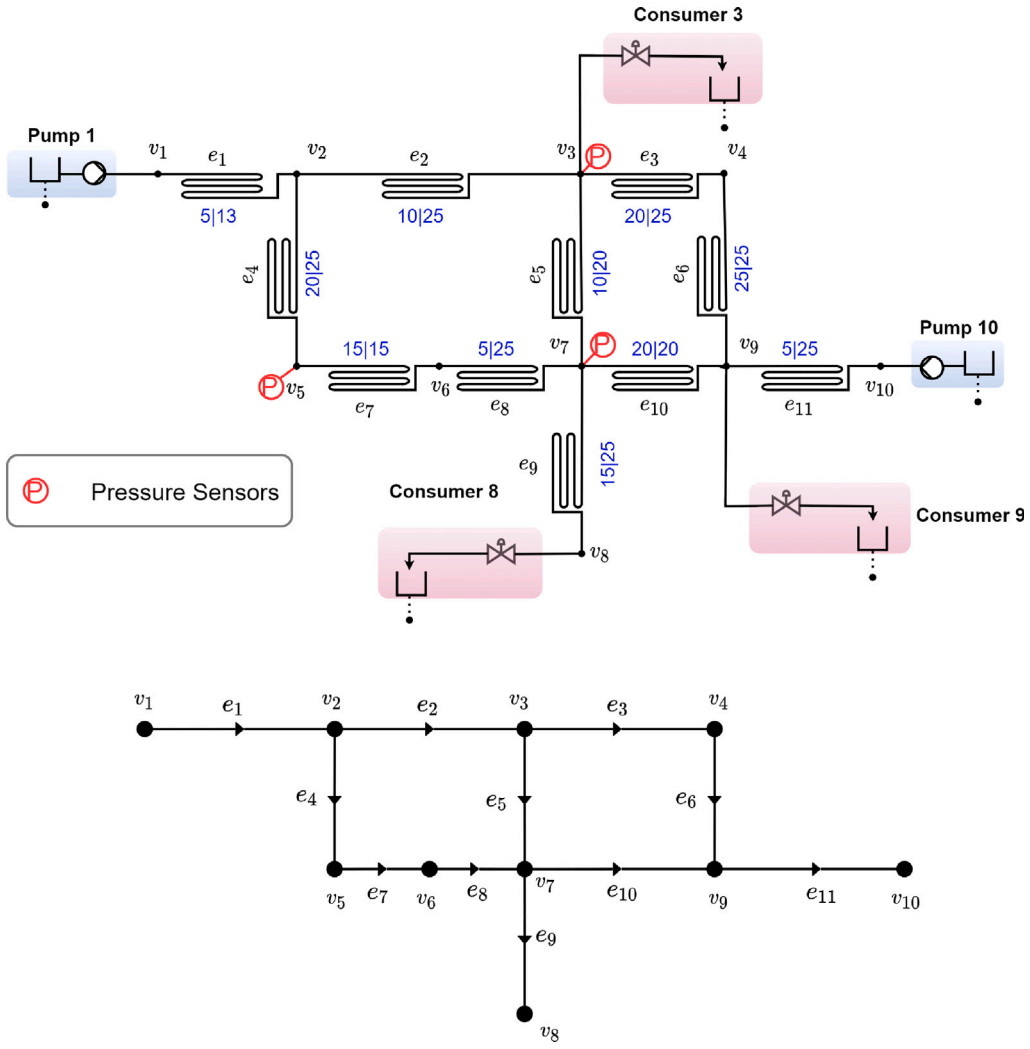

Fig. 4. The SWIL setup and its corresponding network graph used for laboratory test. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where $t_h[k]$ is the time at the k instance and the parameter values are presented in Table 2. These parameters are identified by trial and error to imitate residential consumers, and in the laboratory setup, 120 min operation mimics a full 24 hours of real-life conditions. With the parameters in Table 2, the consumer demand pattern is presented in Fig. 5. Note that as per the convention considered in this work, the consumer demands flow is to have negative values. However, to avoid confusion for the reader, they are represented with their positive values in the figure. Drawing a parallel to a real-life area of residential consumers, in the figure it can be seen that the magnitude of the

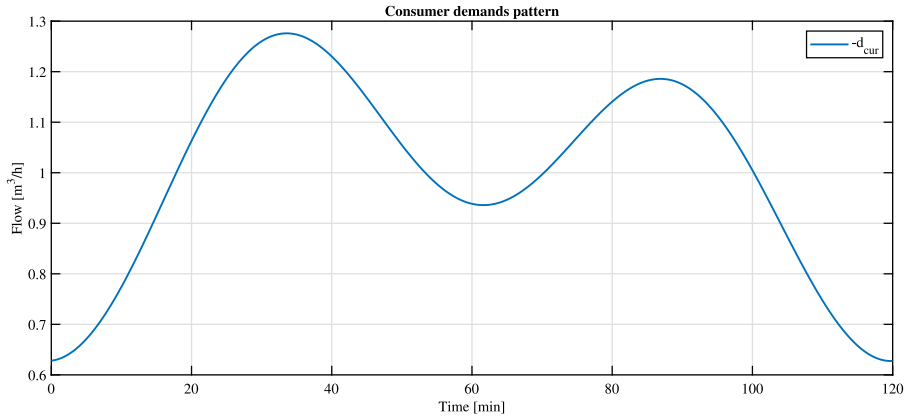

Fig. 5. Consumer demand pattern set-points from (63) and Table 2.

Table 2
Parameters for the consumer demand pattern.

Parameter	Value	Parameter	Value
a_0	−1	ω	0.0522
a_1	0.155	a_2	0.217
b_1	−0.044	b_2	0.005

Table 3
Scaling factor for individual consumer patterns.

Scaling factor	Value
Θ_3	0.18
Θ_8	0.25
Θ_9	0.32

consumer demand is maximum during the morning hours then falls during midday and rises again in the evening before falling to the minimum at night.

From the consumer pattern (63), the set-points for individual consumers can be given as,

$$d_i(t_h[k]) = \Theta_i d_{cur}(t_h[k]), \quad \forall d_i \in d_c, \tag{64}$$

where Θ_i is the scaling factor for individual consumers. The Θ_i for the Consumers 3, 8 and 9, considered in this work are presented in Table 3.

As presented in Section 5, for pressure/flow control of the pumps a local control is to be employed. In this work, a simple Multiple Input Multiple Output State Space (MIMO-SS) controller is used. We use the controller developed by Nielsen, Christensen, and Gnap Simon (2021) which is based on a dynamic model of the laboratory water network setup. Further information about the dynamic model and the design of the MIMO-SS controller can be found in Nielsen et al. (2021).

6.2.1. Laboratory results for leakage detection and localization

Under the nominal condition, the pumps are controlled as per Assumption 1, with Pump 10 maintaining the set pressure and Pump 1 maintaining a fixed distribution of supply flow from it. Fig. 6 presents the flow and pressure measured data for both pumps. In the 1st subplot the set-point for Pump 1 flow, *ref* d_1, the measured Pump 1 flow, d_1, and the measured Pump 10 flow, d_{10}, are presented. In this test, $\kappa = 0.4$ (from (17)) is set and therefore the set-point for Pump 1 flow is calculated as 0.6 times the total supply flow. A leakage is introduced in the network at 5117 s, i.e. around 85 min, which is indicated by *Leak time* in the plots. The 2nd subplot presents the set-point for Pump 10 pressure, *ref* p_{10}, the measured Pump 10 pressure, p_{10}, and the measured Pump 1 pressure, p_1. Here, the set-point for Pump 10 pressure is fixed at 0.4 bar.

Table 4
Identified parameters of the network.

Node	α	β
3	0.0020	−0.0742
5	0.0283	−0.0024
7	−0.0113	−0.0048

Based on (64) and Table 3, the consumer demand set-points are generated which is presented in the first 3 subplots of Fig. 7. The set-points for Consumer 3, 8 and 9 are labelled as *ref* d_3, *ref* d_8 and *ref* d_9 respectively. Fig. 7 also presents the measured consumer demands for these consumers which the PI controller tries to maintain around the set-points. The PI controllers are intentionally not tuned for individual valve control, which introduces noise to the consumer demands in the laboratory test. However, this uncertainty is also present in real-life scenarios and helps to more effectively evaluate the performance of the framework. The 4th subplot presents the leakage flow. The leakage is generated by simply opening a valve at node υ_2 at a fixed degree and letting a free flow of water out of the network. This results in a leakage flow of around 0.22 m^3/h, labelled as δd_2, starting from 85.28 min. As previously mentioned, the convention followed in this work is to consider consumer demand and leakage flow as negative values. Nevertheless, to prevent confusion for the reader, they are illustrated using their positive values in the figures. Furthermore, this leakage size is considered a large leakage compared to the consumer flows. Therefore, it is anticipated that the contamination risk is high.

The measured pressure data of the network is presented in 1st subplot of Fig. 8. As presented before the pressure in the network is measured at nodes υ_3, υ_5 and υ_7, which are labelled in the figure as p_3, p_5 and p_7 respectively. The network parameters α_i and β_i, from (36), are identified for these nodes using these pressure measurements. The identified parameters are listed in Table 4. Further, the residuals are generated from these pressure measurements using (37). These pressure residuals are presented in 2nd, 3rd and 4th subplots of Fig. 8. In the figure, it can be seen that the mean value of residual deviates from zero when the leakage is introduced in the network.

To detect this change in mean we use the GLR decision function from (38). The GLR test results are presented in Fig. 9. The mean value of the residuals under nominal conditions is assumed to be 0. The parameters of GLR, $M_w = 300$ and $thr = 300$ for the residuals are selected based on experimental tests. In the figure, the GLR decision functions for the three residuals r_3, r_5 and r_7 are labelled ϕ_3, ϕ_5 and ϕ_7 respectively. It can be seen that before the leakage is introduced, all three GLR decision functions are well below the threshold. When the leakage is introduced the GLR decision functions rise, and ϕ_3 and ϕ_5

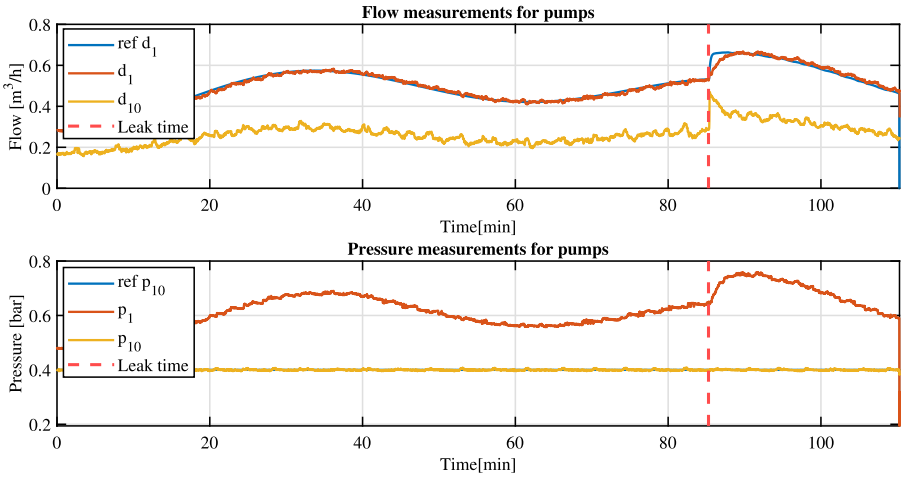

Fig. 6. Flow and pressure measurements for the pumps with their respective reference set-points.

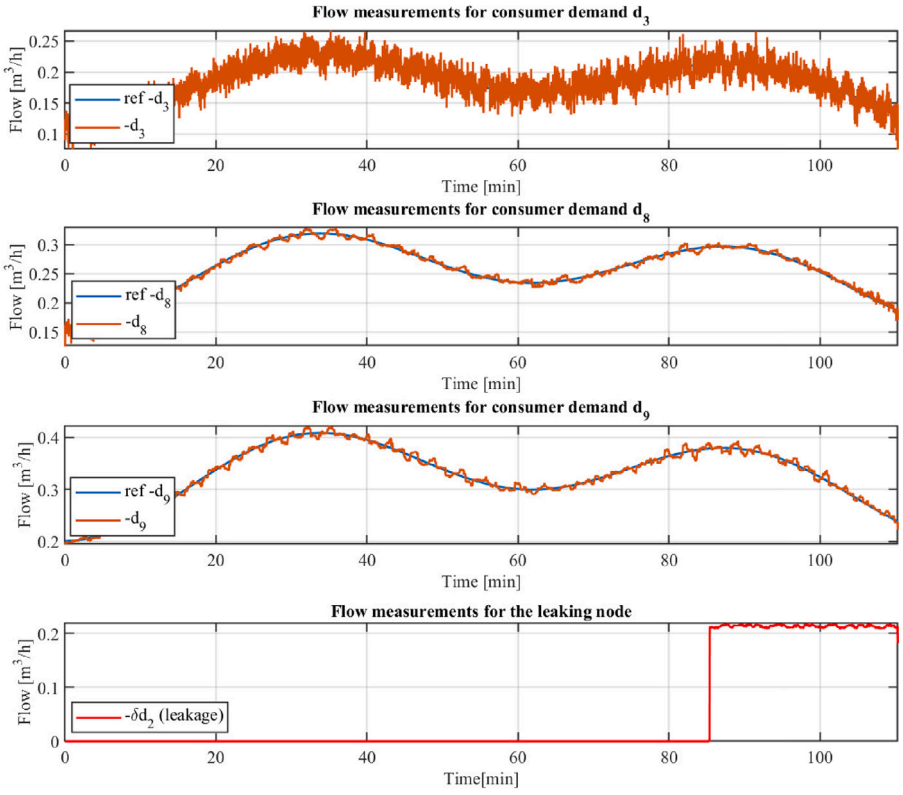

Fig. 7. Measured consumer demand flows and the leakage flow in the laboratory test.

cross the threshold. A leakage alarm is generated as soon as any one of them crosses the threshold. With the actual leakage time being 5117 s, ϕ_5 crosses the threshold at 5127 s Therefore, soon after the leakage is introduced the leakage alarm is generated.

Now, the generated residuals, in Fig. 8, are compared to the residual signature as (43) and generate leakage indicators using (45). The sensitivity matrix, G, required for the comparison is obtained from the network model parameters and the estimate of the distribution of the nominal non-reference flows, W. The leakage indicators from the laboratory test are presented in Fig. 10. In the figure, the likelihood of leakage at a node in the laboratory setup is represented by a colour scale. The colour closest to 1 indicates the highest likelihood of leakage, while the colour closest to 0 indicates the lowest likelihood of leakage. The leak was introduced at node v_2 and in the figure, it can be seen that the leakage indicators also point towards nodes v_2 and v_3 for a higher likelihood of a leak.

Remark 4. The laboratory test presented here utilizes a model for leakage detection and localization that is not highly precise. The specific values for parameters related to the laboratory setup, such as pipe roughness or sensor elevation, are either unknown or not considered in the model. This is similar to a real-life water network, in which the utility may have a general model of the network structure but may not have precise knowledge of all the parameters.

6.2.2. *Laboratory results for contamination mitigation control*

The leakage localization algorithm point indicates the nodes v_2 and v_3 as the leaking nodes. Now, using (49) and (52) the cut edges and the

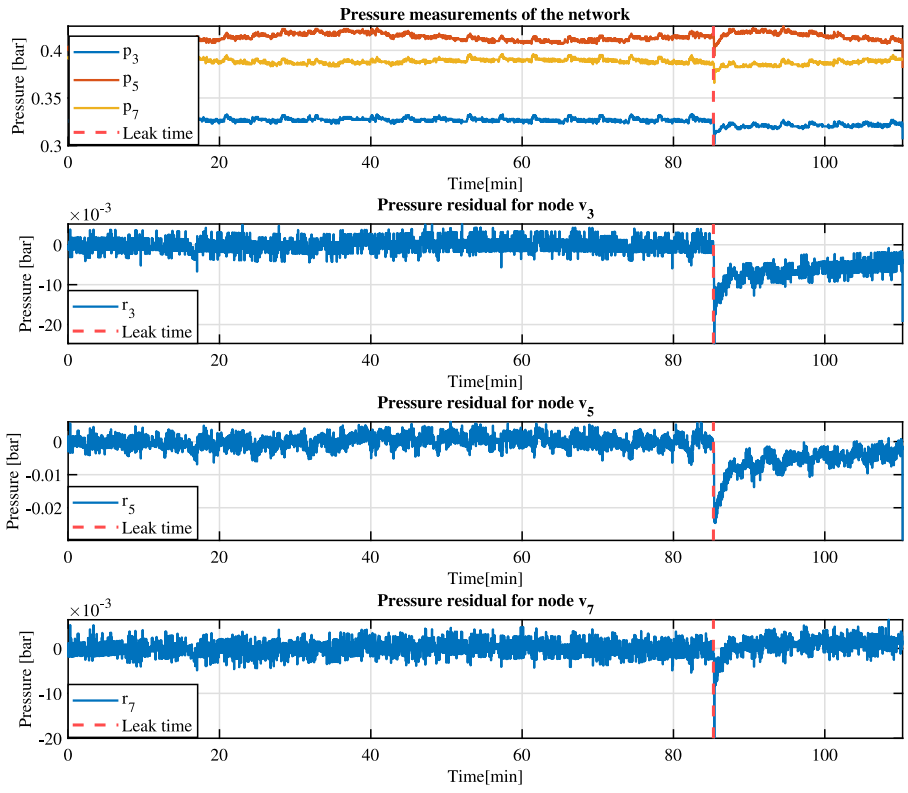

Fig. 8. Pressure measurements of the network and the pressure residuals generated from the laboratory test.

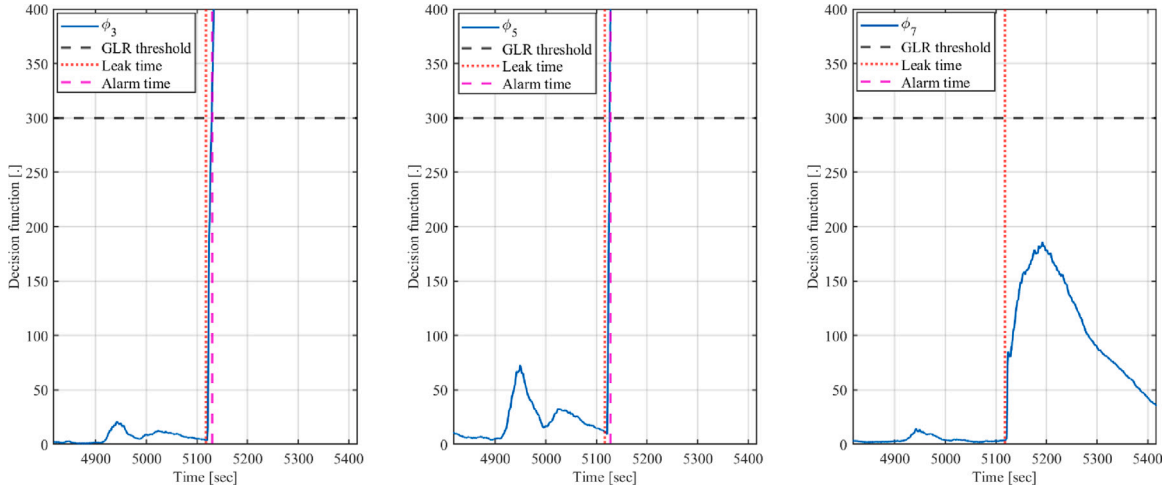

Fig. 9. The GLR test results to detect a change in mean for the pressure residuals presented in Fig. 8.

flow direction in them can be identified. The cut edges, which are e_1, e_3, e_4 and e_5, are coloured green or red in Fig. 10. The green-coloured edges are supposed to have the flow direction the same as the direction of the graph, the red-coloured edges are supposed to have the flow direction opposite to the direction in the graph and for the rest of the edges in the black, there is no constraint on the flow direction. These flow directions are used as constraints in the contamination mitigation control problem.

After the leakage alarm is generated, the contamination mitigation control comes is activated. The flow set-point for Pump1 and the pressure set-point for Pump 10 are obtained from solving the optimization problem (60) subject to (61) and (62). The optimization problem was solved using IPOPT solver (Wächter & Biegler, 2006) which was interfaced by CasADi (Andersson, Gillis, Horn, Rawlings, & Diehl, 2019). In the laboratory test, the critical nodes, at which the pressure is to be at the desired level, are set as the consumer nodes v_3, v_8 and v_9. The other parameters of the optimization problem are listed in Table 5. The results from the contamination mitigation control part are presented in Figs. 11–13.

Table 5
Parameters for the optimization problem.

Parameter	Value	Parameter	Value
p_c^*	0.3 bar	p_s^{max}	0.7 bar
p_c^{min}	0.2 bar	d_s^{max}	2 m^3/h
p_c^{max}	0.4 bar		

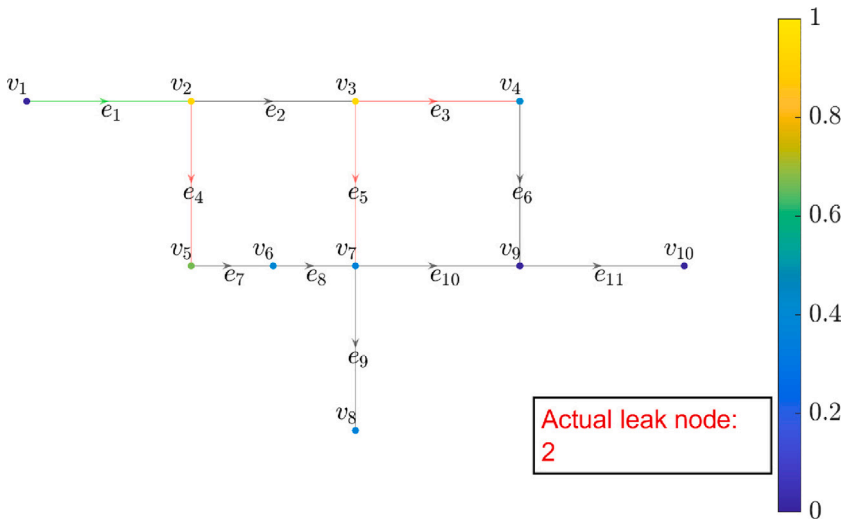

Fig. 10. The leakage localization indicators for each node along with cut edge flow direction indicators from the laboratory test. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

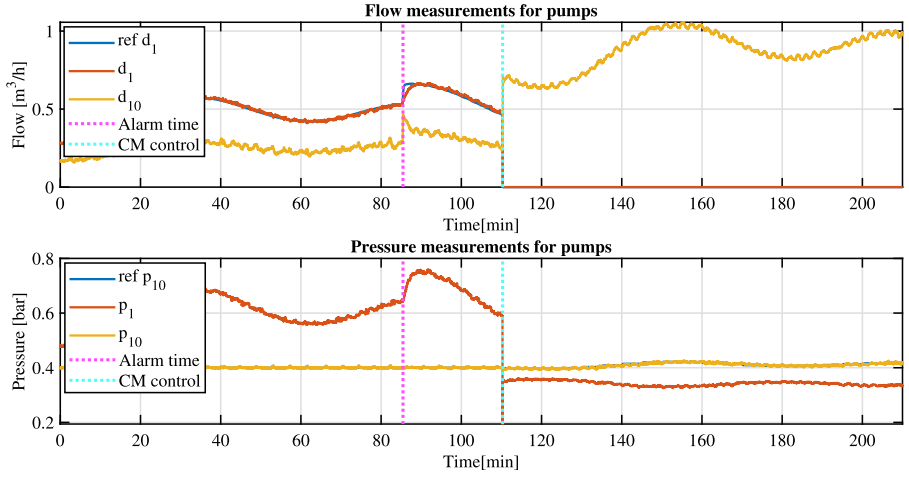

Fig. 11. Flow and pressure measurements for the pumps with their respective reference set-points after the contamination mitigation control is activated.

Fig. 11 presents the flow and pressure measurements for both pumps. In the figure, *CM control* represents the time when the contamination mitigation control takes over, after which the flow set-point no longer follows Assumption 1. The flow and pressure set-points are given by the contamination mitigation control and the local controller regulates the pump speeds to maintain desired set-points.

As mentioned before, the critical nodes considered in this work are the consumer nodes and the pressure measurements for the consumer nodes are presented in Fig. 12. The contamination mitigation control aims to keep the consumer pressure near the target value of 0.3 bar, but it must also consider the constraints of the optimization problem.

Finally, Fig. 13 presents the flow measurements in the cut edges of the network. As previously mentioned, the cut edges are e_1, e_3, e_4 and e_5, and in the figure, the flow in these edges are labelled as q_1, q_3, q_4 and q_5. From Fig. 10, the flow direction in e_1 should be the same as the graph, i.e. q_1 should be positive or zero. And from 1st subplot of Fig. 13 it can be seen that q_1 is already positive but after the contamination mitigation control is activated it becomes zero. Again, from Fig. 10, the flow direction in e_3, e_4 and e_5 should be the opposite of the graph, i.e. q_3, q_4 and q_5 should be negative or zero. And from the 2nd subplot of Fig. 13 it can be seen that q_3 is positive before and as soon as the contamination mitigation control is activated it changes to negative. From subplots 3 and 4, q_4 and q_5 are positive but go to zero after the contamination mitigation control is activated. The flow constraint for q_4 and q_5 are satisfied, as per (62a), but they do not go negative and similarly q_1 does not stay positive, this could be due to other constraints in the optimization problem and this being the best solution. Note that the flow in q_1, q_4 and q_5 seems to be exactly zero in the figure, however, there could be a minor flow that is measured as zero due to the sensor's lower range limit.

6.3. L-town benchmark network

To evaluate the framework's performance on a large-scale network, a test is conducted on an EPANET model of the water network referred to as L-town. This network is based on a coastal city in Cyprus and is presented as a benchmark water distribution network by Vrachimis et al. (2022).

In this work, only Area A of the L-town network is considered, and the EPANET model (Rossman et al., 2000) of the network is presented in Fig. 14. The network comprises 657 nodes and 762 pipes or edges, with a loop ratio of 25%, which is a measure of network complexity (Vrachimis, Timotheou, Eliades, & Polycarpou, 2019). There are two supply nodes, *Supply 1* and *Supply 2*, which are marked with pink squares in Fig. 14. There are two different types of consumer patterns: residential and commercial, and each node has different consumer demands, which is a linear combination of the base demands with the

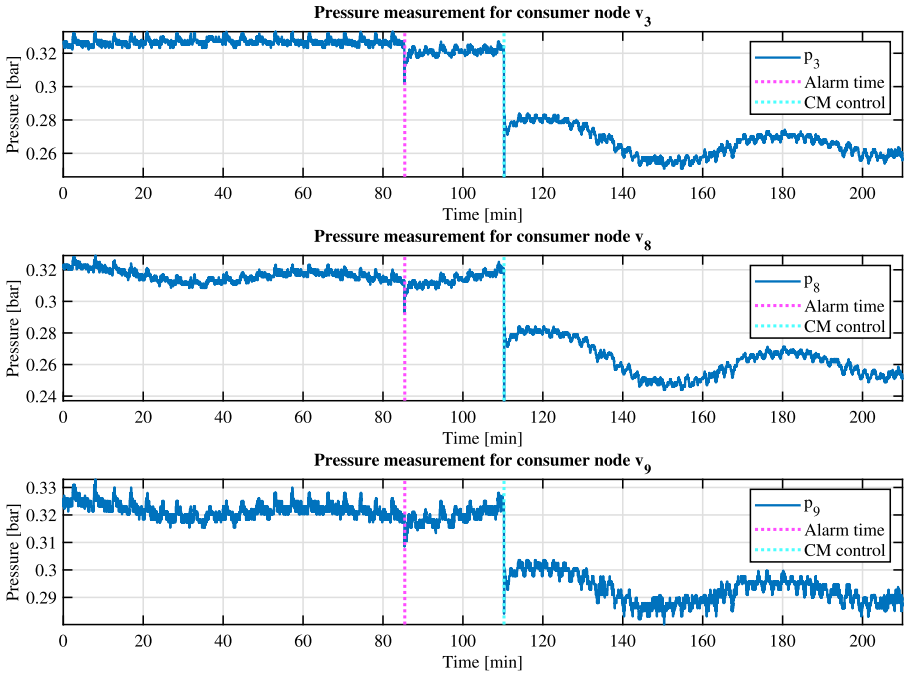

Fig. 12. Pressure measurements at the critical nodes under the contamination mitigation control.

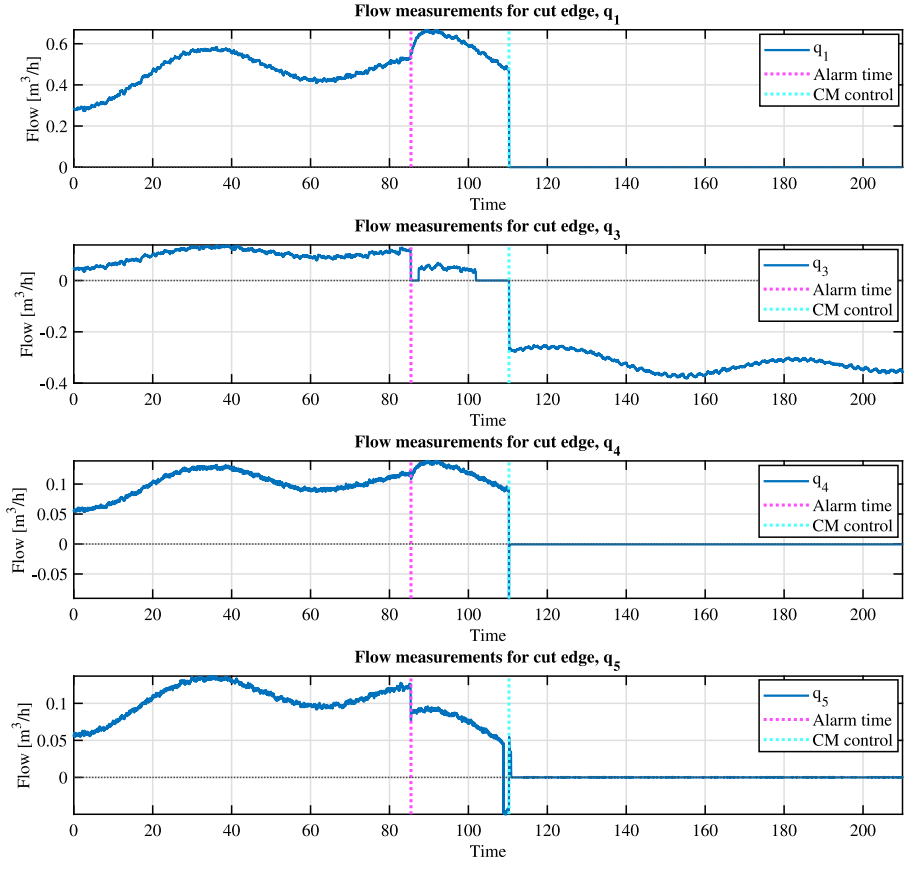

Fig. 13. Flow measurements in the cut edges of the network under the contamination mitigation control.

corresponding patterns. In this work, we measure the network pressure at 13 nodes, marked with a red star in Fig. 14. The number next to the stars is the node number from the EPANET model. For the study, a leakage is generated at node *n174*, marked with a brown diamond. The conducted test and its results are presented in the following section.

6.4. L-town network test results

The L-town water network is simulated on EPANET and the data is generated with a sampling time of 10 minutes. In the test, under both the nominal control and the contamination mitigation control, *Supply*

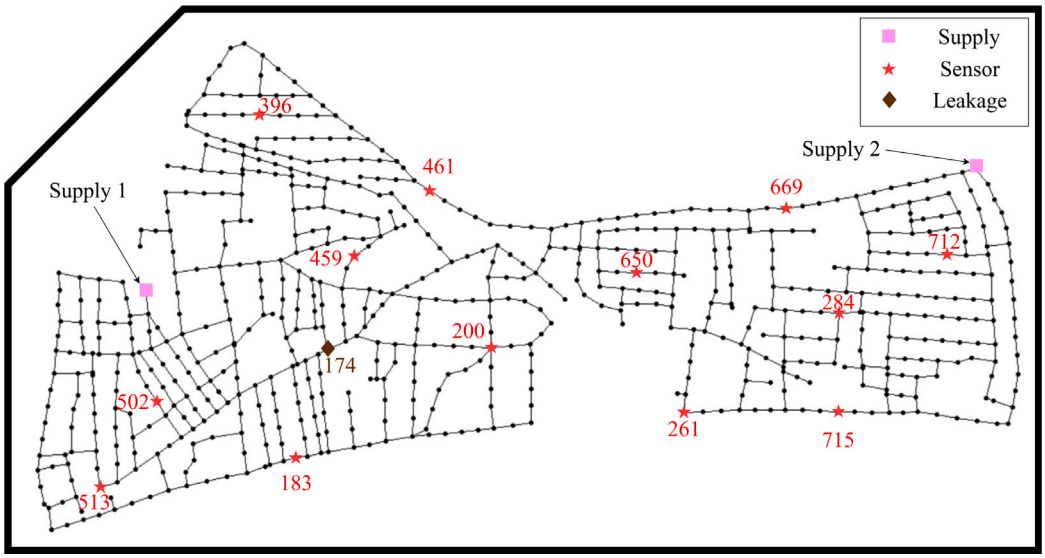

Fig. 14. The area A of the L-town benchmark network (Vrachimis et al., 2022) with the pressure sensor placement considered in the work.

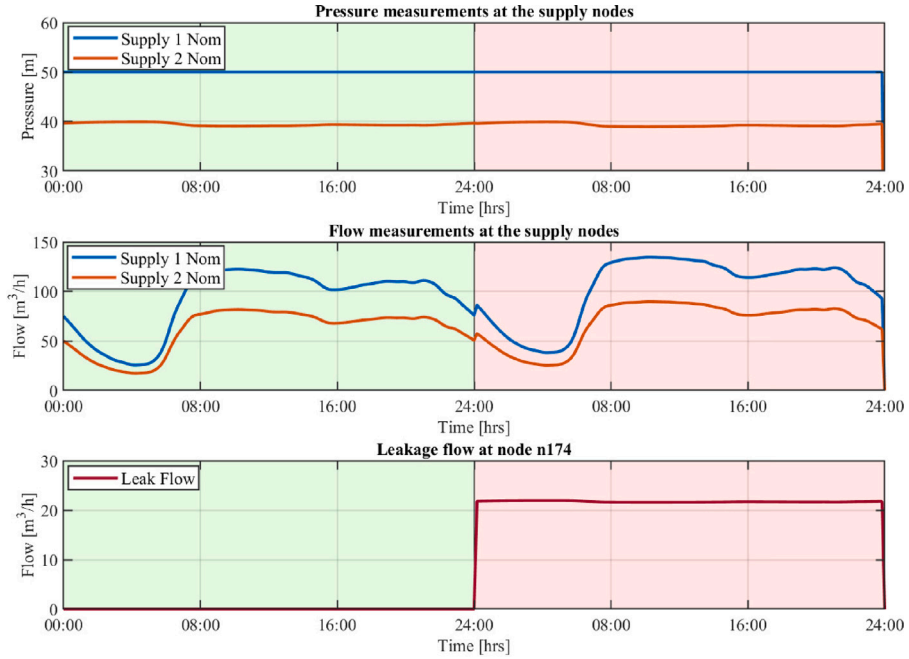

Fig. 15. Pressure and flow data of the supply nodes, and the leakage flow with the nominal controller. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

1 is pressure controlled, while *Supply 2* is flow controlled. Further, the leakage diagnosis and contamination mitigation test results for the L-town network are presented.

6.4.1. Leakage diagnosis test results for the L-town network

The 1st and 2nd subplots of Fig. 15 present the pressure and flow of the supply nodes. A leakage is generated at the end of the first day, and the green shaded area represents the network condition with no leakage in the test, while the red shaded area represents the condition with a leakage in the network. This figure presents the results in which the supply nodes are controlled by the nominal controller, as per Assumption 1, throughout the test. Under nominal control, the pressure at the *Supply 1* node is maintained at 50 m, and the flow at the *Supply 2* node is maintained at 0.4 times the total supply flow. The leakage at node *n174* is generated using the emitter equation (Rossman et al., 2000) by setting the emitter coefficient in EPANET, resulting in a leakage flow of approximately 21.5 m^3/h in the network. This can also be seen with the slight increase in the supply flow at 24:00 hours on the first day.

The pressure in the network is measured at 13 nodes, as shown in Fig. 14. Using these pressure measurements, the network parameters α_i and β_i are identified for these 13 nodes using Eq. (36). Additionally, pressure residuals are generated using this model and Eq. (37), and the resulting pressure residuals are presented in Fig. 16. When leakage is introduced into the network at 24:00 hours, the mean value of the residuals deviates from zero. In the no-leakage case, the residuals also deviate slightly from zero just before 08:00 hours. This can be explained by the fact that not all consumers follow the same demand pattern which is an assumption in the development of this leakage diagnosis framework. Again, methods such as GLR can be used to detect changes in the mean value of the residual.

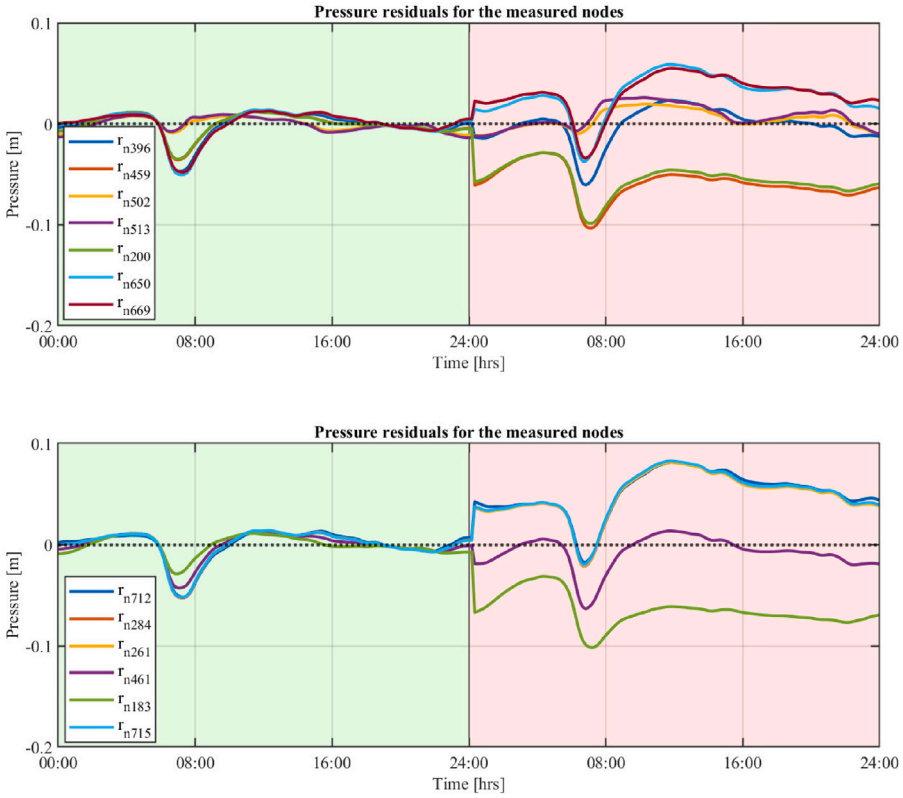

Fig. 16. Pressure residuals generated for the measured nodes from the L-town network test.

Furthermore, the leakage indicator for each node is computed based on the generated residual by comparing it to the residual signature using (45) and (43). Since not all consumers follow the same demand pattern, the distribution of the nominal non-reference nodal flows, W, is not constant. Therefore, the sensitivity matrix, G, is computed with the mean value of W. The leakage indicators obtained from the L-town network test are presented in Fig. 17. The likelihood of leakage at a node is represented by a colour scale, where the colour closest to 1 indicates the highest likelihood of leakage, and the colour closest to 0 indicates the lowest likelihood of leakage. The actual leakage node is marked with a brown diamond in the figure. From the indicators, it is evident that the leakage localization algorithm can localize the leakage to an area around the actual leakage point. However, with this low number of sensors compared to the total number of nodes, it is not possible to detect the actual leaking node, as the residual signatures would be nearly identical for a set of nodes close to each other. Note that even with multiple types of consumers in the network, the leakage detection and localization framework produces satisfactory results.

Nodes with a leakage indicator value higher than 0.9 are considered part of the leakage node set. This threshold is chosen based on experimental results. Based on this, the cut edges and their flow directions are identified using (49) and (52). These cut edges are highlighted in Fig. 17 with a neon-green marker. Eight cut edges are identified, and from the figure, it can be seen that these edges isolate the leaking nodes from the rest of the network. The flow in the edges on the right of the leakage node set is to be from right to left, and the flow in the edges on the left of the leakage node set is to be from left to right. The required flow directions in the cut edges with respect to the edge direction in the digraph of the network are provided in the E matrix from (49). This information is included as a constraint in the contamination mitigation control problem using (52).

6.4.2. Contamination mitigation control test results for the L-town network

After detecting and localizing the leakage, the contamination mitigation control is activated. The contamination mitigation control then controls the pressure at *Supply 1* and the flow ratio at *Supply 2*. To obtain pressure and flow ratio set-points, the control solves the optimization problem (60) subject to (61) and (62). The optimization problem is solved using the IPOPT solver (Wächter & Biegler, 2006) which is interfaced with CasADi (Andersson et al., 2019). In the L-town network test, all non-supply nodes are selected as critical nodes. A pressure of 35 m is desired to be maintained at these nodes with a minimum pressure of 25 m and a maximum pressure of 55 m. The pressure and flow ratio set-points are given to EPANET to simulate the water network. As previously mentioned, this is done at a sampling time of 10 minutes.

Figs. 18, 19, and 20 present a comparison of the network behaviour under nominal control and contamination mitigation control. Subplots 1 and 2 of Fig. 18 compare the pressure and flow from the supply nodes between nominal control and contamination mitigation control. The network pressure and flow with nominal control are presented with solid lines, while the same with contamination mitigation control are presented with dashed lines. Nominal control maintains a pressure of 50 m at *Supply 1* and a flow ratio of 0.4 at *Supply 2*, even under leakage conditions. However, contamination mitigation control operates the network at a slightly lower pressure, with a pressure of 45 m at *Supply 1*, which optimizes the objective of achieving a desired pressure of 35 m at all non-supply nodes. Note that contamination mitigation control is only activated during leakage conditions. Under contamination mitigation control, the flow ratio at *Supply 2* ranges between 0.45 and 0.6 of the total flow. These results are expected because the leakage is on the left side of the network, and *Supply 2* now needs to supply a higher flow ratio to direct the flow towards the leakage node set. The third subplot of Fig. 18 compares the leakage flows under nominal control and contamination mitigation control. There is only a slight difference between the leakage flows, and even with directing the flows towards the leakage area under contamination mitigation control, the leakage magnitude does not increase.

Fig. 19 presents a comparison of the cut edge flows in edges where the flow needs to be positive under leakage conditions to mitigate the

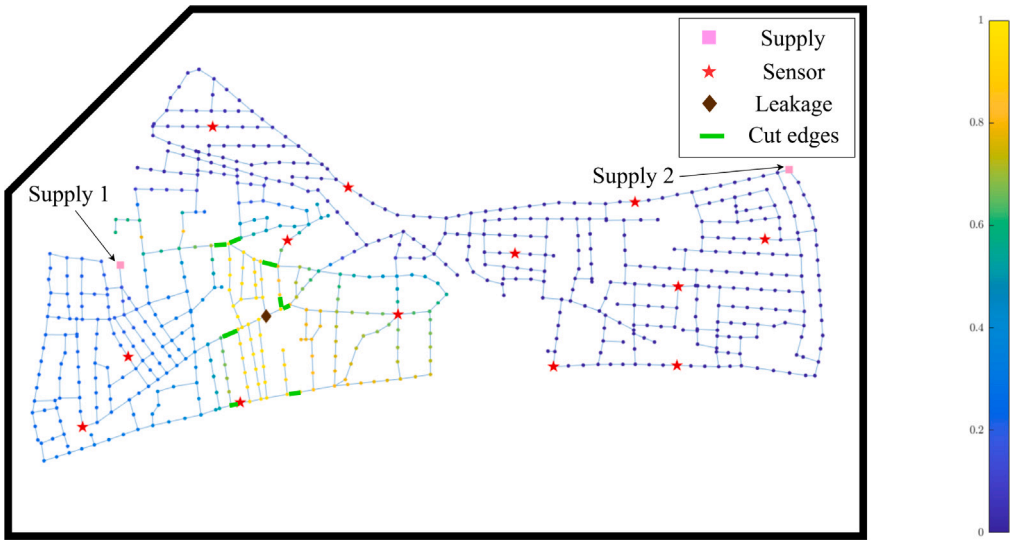

Fig. 17. The leakage localization indicators for each node along with marked cut edges from the L-town network test. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

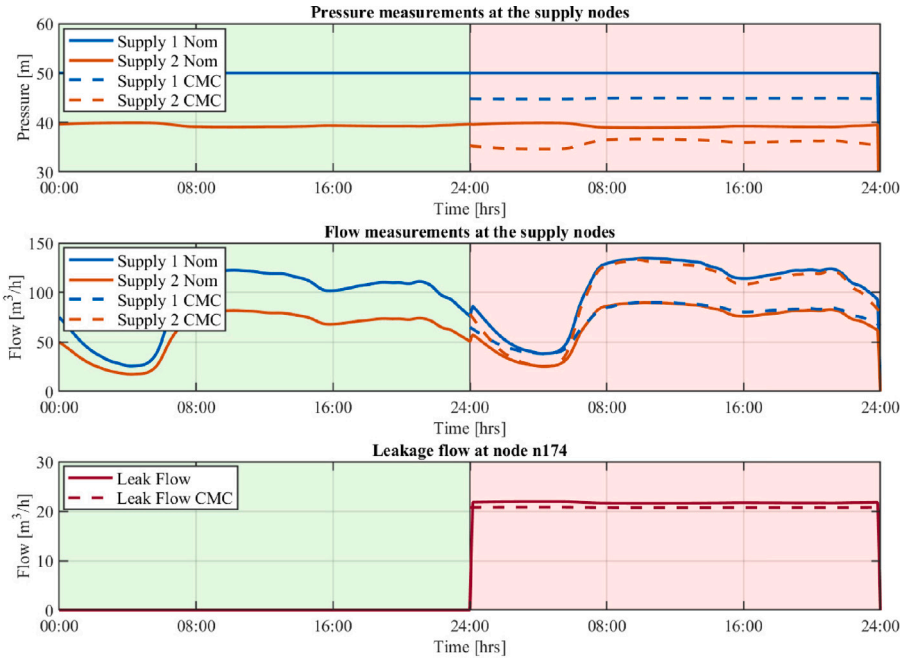

Fig. 18. A comparison of pressure and flow data of the supply nodes, and the leakage flow between the network controlled with nominal control and contamination mitigation control. Solid lines represent the network pressure and flow under nominal control, whereas dashed lines represent the network pressure and flow under contamination mitigation control.

spread of contaminants, between nominal control and contamination mitigation control. Again, the positive value constraint is with respect to the edge direction in the network digraph. In the figure, it can be seen that with nominal control, the flow is always positive in edges *p117*, *p202*, and *p204*. However, under leakage conditions, the flow is negative for some duration in edge *p595*. With contamination mitigation control, the magnitude of the flow is decreased, but it is always positive.

Fig. 20 presents a comparison of cut edge flows in the edges, where the flow needs to be negative under leakage conditions, between the network controlled with nominal control and contamination mitigation control. Under nominal control, the flow in edges *p127*, *p231*, and *p629* is mostly positive under leakage conditions. However, under contamination mitigation control, the flow in these edges is mostly negative. The flow constraint in edge *p629* is not violated in the initial duration just after the leakage. However, a non-zero value for the slack variable, ϵ from (62a), is not observed, which could be due to the unknown leakage flow that is not taken into account in the optimization problem. The flow in edge *p136* is mostly negative under nominal control and always negative with contamination mitigation control.

7. Limitations and future work

This section discusses the limitations of the work and potential areas for future research. The limitations for each section — graph theory based modelling, leakage diagnosis, and contamination mitigation control — are discussed separately since they can be utilized independently for different applications. For instance, if the leakage localization algorithm presented in this work is not applicable to a certain network,

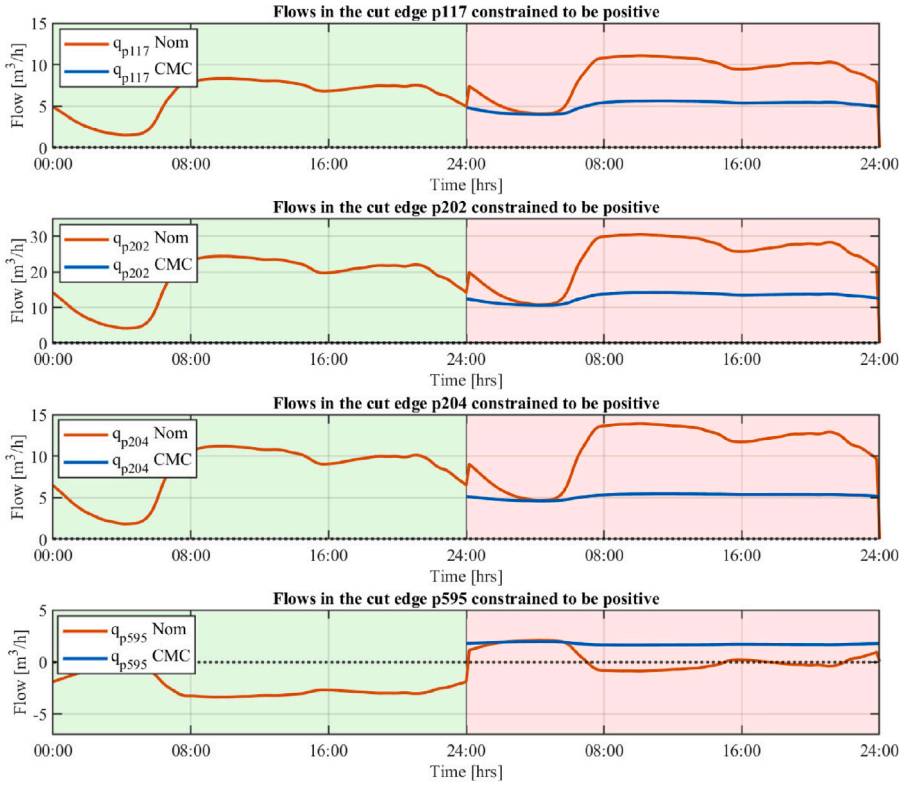

Fig. 19. A comparison of cut edge flows in the edges, where the flow is required to be positive under leakage conditions, between the network controlled with nominal control and the contamination mitigation control. The legend *Nom* at the end represents the flow under nominal control, whereas the legend *CMC* represents the flow under contamination mitigation control.

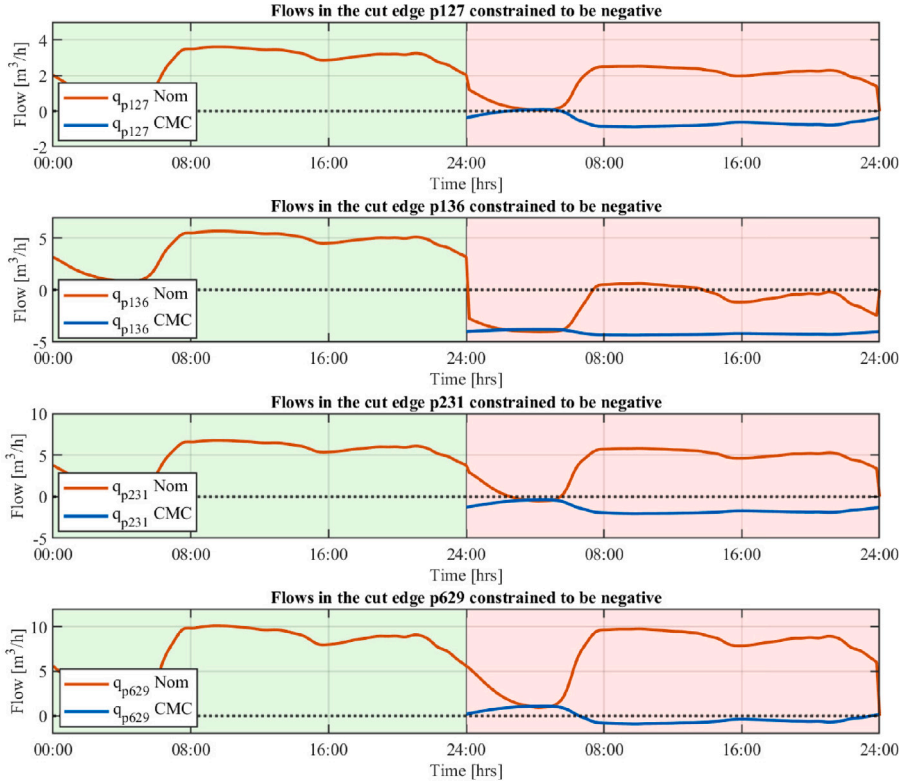

Fig. 20. A comparison of cut edge flows in the edges, where the flow is required to be negative under leakage conditions, between the network controlled with nominal control and the contamination mitigation control. The legend *Nom* at the end represents the flow under nominal control, whereas the legend *CMC* represents the flow under contamination mitigation control.

a different localization scheme can be used in conjunction with the contamination mitigation control framework.

- **Graph theory based modelling:** The graph theory based model presented in this paper is utilized to develop frameworks for leakage localization and contamination mitigation control. In these applications, the network is assumed to be in a steady state condition, and thus the dynamics of the network flow are not presented here. Additionally, since the leakage localization algorithm is not compatible for networks with tanks, the tank dynamic model is also excluded. However, for different applications, readers can refer to Rathore (2020), which presents the graph theory based model with the flow and tank dynamics. Moreover, while the demand flows in this work are modelled as pressure independent, they can be modelled as pressure dependent by considering the flow dynamics. Furthermore, Rathore (2020) also presents models for other network components such as valves and pumps.
- **Leakage diagnosis:** The leakage diagnosis algorithm is developed based on a reduced order model that assumes only one type of consumer in a DMA with a fixed distribution of nodal flows. However, this is a strong assumption and limits the applicability of the algorithm to only certain types of networks. To overcome this limitation, a proposed solution for DMAs with more than one type of consumer is to divide the network's operational day into multiple time segments, where the distribution is approximately constant within each segment, and analyse them separately. Further research and testing would be required to validate this approach. Furthermore, the current diagnosis algorithm is not compatible with networks that have tanks. This is due to the fact that the tank flow changes direction during the network operation. If the supply flow is higher than the consumer demands, water flows into the tank, and acts as a consumer. Conversely, when the consumer demands are higher than the supply flow, the tank supplies the network and acts as a supply point. As a result, the distribution of nodal flows will not be constant and could be difficult to predict. Improving the leakage diagnosis algorithm to account for tank behaviour is seen as a potential avenue for future work.
- **Contamination mitigation control:** The contamination mitigation control developed in this study utilizes parameters based on the known laboratory setup, which are approximate values, and hence are not highly accurate. The uncertainties in the model parameters were not taken into account during the development of the control, which could result in a less-than-ideal outcome compared to the expected result. It is worth noting that this limitation would also be present in real-life scenarios, where the control design is based on a standard hydraulic model, such as EPANET, which is also typically not highly accurate. Additionally, the computation of optimal supply pressure and flow is based on expected consumer demands, and uncertainties in these values could also lead to less-than-ideal results. Moreover, the network's actuators in this study were limited to pumps, but including shut-off valves or pressure/flow control valves in the network could help direct the flow towards the leakage area. This is seen as an area of future work. Another potential contribution could be formulating the optimization problem using pressure gradients in the network to accomplish the same goal. Additionally, minimizing water loss while maintaining pressure could also be added as an objective to the optimization problem.

8. Conclusion

Overall, this work aims to improve the efficiency of localizing leakages and, address and prevent contamination due to these leakages. This paper presents an extension of previous research (Rathore et al., 2022) on leakage detection and localization to include contamination mitigation control. A data-driven model is used to generate pressure residuals based on which the probable location of leakage is identified. A contamination mitigation control is designed to prevent contamination from occurring due to leakage. The control is formulated as an optimization problem with a constraint to direct the flow of water in the network towards the leakage area. This constraint is formulated using a novel method that identifies the pipes and their flow direction in them based on the leakage localization information. Finally, the combined leakage diagnosis and contamination mitigation control framework is tested on a laboratory water network setup and a large-scale benchmark water network, L-town.

Declaration of competing interest

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

Data availability

Data will be made available on request

References

Adedeji, K. B., Hamam, Y., Abe, B. T., & Abu-Mahfouz, A. M. (2017). Towards achieving a reliable leakage detection and localization algorithm for application in water piping networks: An overview. *IEEE Access*, *5*, 20272–20285.

Alfonso, L., Jonoski, A., & Solomatine, D. (2010). Multiobjective optimization of operational responses for contaminant flushing in water distribution networks. *Journal of Water Resources Planning and Management*, *136*(1), 48–58.

Andersson, J. A. E., Gillis, J., Horn, G., Rawlings, J. B., & Diehl, M. (2019). CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, *11*(1), 1–36. http://dx.doi.org/10.1007/s12532-018-0139-4.

ASCE (2004). *Interim voluntary guidelines for designing an online contaminant monitoring system*. Reston.

Bazargan-Lari, M. R., Taghipour, S., & Habibi, M. (2021). Real-time contamination zoning in water distribution networks for contamination emergencies: A case study. *Environmental Monitoring and Assessment*, *193*(6), 336.

Besner, M. -C., Ebacher, G., Jung, B., Karney, B., Lavoie, J., Payment, P., et al. (2010). Negative pressures in full-scale distribution system: Field investigation, modelling, estimation of intrusion volumes and risk for public health. *Drinking Water Engineering and Science*, *3*(2), 101–106.

Besner, M. -C., Prévost, M., & Regli, S. (2011). Assessing the public health risk of microbial intrusion events in distribution systems: Conceptual model, available data, and challenges. *Water Research*, *45*(3), 961–979.

Blanke, M., Kinnaert, M., Lunze, J., Staroswiecki, M., & Schröder, J. (2006). *vol. 2, Diagnosis and fault-tolerant control*. Springer.

Casillas, M. V., Garza-Castañón, L. E., & Puig, V. (2013). Extended-horizon analysis of pressure sensitivities for leak detection in water distribution networks: Application to the Barcelona network. In *2013 european control conference* (pp. 401–409). IEEE.

Casillas, M. V., Garza-Castañón, L. E., Puig, V., & Vargas-Martinez, A. (2015). Leak signature space: An original representation for robust leak location in water distribution networks. *Water*, *7*(3), 1129–1148.

Casillas Ponce, M. V., Garza Castanon, 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.

Chan, T. K., Chin, C. S., & Zhong, X. (2018). Review of current technologies and proposed intelligent methodologies for water distributed network leakage detection. *IEEE Access*, *6*, 78846–78867.

Deng, Y., Jiang, W., & Sadiq, R. (2011). Modeling contaminant intrusion in water distribution networks: A new similarity-based DST method. *Expert Systems with Applications*, *38*(1), 571–578.

Deo, N. (2017). *Graph theory with applications to engineering and computer science*. Courier Dover Publications.

Di Nardo, A., Di Natale, M., Guida, M., & Musmarra, D. (2013). Water network protection from intentional contamination by sectorization. *Water Resources Management*, *27*, 1837–1850.

Ebacher, G., Besner, M. -C., Prévost, M., & Allard, D. (2010). Negative pressure events in water distribution systems: Public health risk assessment based on transient analysis outputs. In *Water distribution systems analysis 2010* (pp. 471–483).

Eliades, D. G., Kyriakou, M., Vrachimis, S., & Polycarpou, M. M. (2016). EPANET-MATLAB toolkit: An open-source software for interfacing EPANET with MATLAB. In *Proc. 14th international conference on computing and control for the water industry* (p. 8). The Netherlands: http://dx.doi.org/10.5281/zenodo.831493.

Eliades, D., Lambrou, T., Panayiotou, C. G., & Polycarpou, M. M. (2014). Contamination event detection in water distribution systems using a model-based approach. *Procedia Engineering*, *89*, 1089–1096.

Eliades, D. G., Stavrou, D., Vrachimis, S. G., Panayiotou, C. G., & Polycarpou, M. M. (2015). Contamination event detection using multi-level thresholds. *Procedia Engineering*, *119*, 1429–1438.

Fontanazza, C. M., Notaro, V., Puleo, V., Nicolosi, P., & Freni, G. (2015). Contaminant intrusion through leaks in water distribution system: Experimental analysis. *Procedia Engineering*, *119*, 426–433.

Guidorzi, M., Franchini, M., & Alvisi, S. (2009). A multi-objective approach for detecting and responding to accidental and intentional contamination events in water distribution systems. *Urban Water Journal*, *6*(2), 115–135.

Hamilton, S., & Charalambous, B. (2013). *Leak detection: Technology and implementation*. IWA Publishing.

Hu, C., Wang, Q., Gong, W., & Yan, X. (2022). Multi-objective deep reinforcement learning for emergency scheduling in a water distribution network. *Memetic Computing*, *14*(2), 211–223.

Islam, N., Farahat, A., Al-Zahrani, M. A. M., Rodriguez, M. J., & Sadiq, R. (2015). Contaminant intrusion in water distribution networks: Review and proposal of an integrated model for decision making. *Environmental Reviews*, *23*(3), 337–352.

Kallesøe, C. S., Jensen, T. N., & Wisniewski, R. (2015). Adaptive reference control for pressure management in water networks. In *2015 european control conference* (pp. 3268–3273). IEEE.

Karim, M. R., Abbaszadegan, M., & LeChevallier, M. (2003). Potential for pathogen intrusion during pressure transients. *Journal American Water Works Association*, *95*(5), 134–146.

Kirmeyer, G. J. (2001). *Pathogen intrusion into the distribution system*. American Water Works Association.

Lambrou, T. P., Anastasiou, C. C., Panayiotou, C. G., & Polycarpou, M. M. (2014). A low-cost sensor network for real-time monitoring and contamination detection in drinking water distribution systems. *IEEE Sensors Journal*, *14*(8), 2765–2772.

Ledesma, J. V., Wisniewski, R., & Kallesøe, C. S. (2021). Smart water infrastructures laboratory: Reconfigurable test-beds for research in water infrastructures management. *Water*, *13*(13), 1875.

Liemberger, R., & Wyatt, A. (2019). Quantifying the global non-revenue water problem. *Water Supply*, *19*(3), 831–837.

Madsen, H. (2007). *Time series analysis*. CRC Press.

Mazzolani, G., Berardi, L., Laucelli, D., Simone, A., Martino, R., & Giustolisi, O. (2017). Estimating leakages in water distribution networks based only on inlet flow data. *Journal of Water Resources Planning and Management*, *143*(6), Article 04017014.

Moghaddam, A., Afsharnia, M., & Peirovi Minaee, R. (2020). Preparing the optimal emergency response protocols by MOPSO for a real-world water distribution network. *Environmental Science and Pollution Research*, *27*, 30625–30637.

Mora-Rodríguez, J., Delgado-Galván, X., Ortiz-Medel, J., Ramos, H. M., Fuertes-Miquel, V. S., & López-Jiménez, P. A. (2015). Pathogen intrusion flows in water distribution systems: According to orifice equations. *Journal of Water Supply: Research and Technology—AQUA*, *64*(8), 857–869.

Nielsen, E. H., Christensen, M., & Gnap Simon, N. (2021). *Leakage detection and localization in water distribution networks with multiple inlets*. Aalborg University, Department of Electronic Systems, URL https://projekter.aau.dk/projekter/files/421098402/CA10_Leakage_localization.pdf.

Palleti, V. R., Kurian, V., Narasimhan, S., & Rengaswamy, R. (2018). Actuator network design to mitigate contamination effects in water distribution networks. *Computers & Chemical Engineering*, *108*, 194–205.

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.

Perez, R., Sanz, G., Puig, V., Quevedo, J., Escofet, M. A. C., Nejjari, F., et al. (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 Magazine*, *34*(4), 24–36.

Ponce, M. V. C., Castañón, L. E. G., & Cayuela, V. P. (2012). Extended-horizon analysis of pressure sensitivities for leak detection in water distribution networks. *IFAC Proceedings Volumes*, *45*(20), 570–575.

Poulin, A., Mailhot, A., Grondin, P., Delorme, L., Periche, N., & Villeneuve, J. -P. (2008). Heuristic approach for operational response to drinking water contamination. *Journal of Water Resources Planning and Management*, *134*(5), 457–465.

Preis, A., & Ostfeld, A. (2008a). Multiobjective contaminant response modeling for water distribution systems security. *Journal of Hydroinformatics*, *10*(4), 267–274.

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.

Rasekh, A., & Brumbelow, K. (2014). Drinking water distribution systems contamination management to reduce public health impacts and system service interruptions. *Environmental Modelling & Software*, *51*, 12–25.

Rathore, S. S. (2020). *Nonlinear optimal control in water distribution network*. Aalborg University, Department of Electronic Systems, URL https://projekter.aau.dk/projekter/files/334475797/Report_1030_Nonlinear_Optimal_Control.pdf.

Rathore, S. S., Kallesøe, C. S., & Wisniewski, R. (2022). Application of leakage localization framework for water networks with multiple inlets in smart water infrastructures laboratory at AAU. *IFAC-PapersOnLine*, *55*(6), 451–457.

Rathore, S. S., Kallesøe, C. S., Wisniewski, R., & Jensen, T. N. (2021). Leakage localization in municipal water supply using self adaptive reduced network models and sensitivity analysis. In *2021 5th international conference on control and fault-tolerant systems* (pp. 199–204). IEEE.

Rossman, L. A., et al. (2000). *EPANET 2: Users manual*. US Environmental Protection Agency. Office of Research and Development.

Sadiq, R., Kleiner, Y., & Rajani, B. (2006). Estimating risk of contaminant intrusion in water distribution networks using Dempster–Shafer theory of evidence. *Civil Engineering and Environmental Systems*, *23*(3), 129–141.

Shafiee, M. E., & Berglund, E. Z. (2015). Real-time guidance for hydrant flushing using sensor-hydrant decision trees. *Journal of Water Resources Planning and Management*, *141*(6), Article 04014079.

Shafiee, M. E., & Berglund, E. Z. (2017). Complex adaptive systems framework to simulate the performance of hydrant flushing rules and broadcasts during a water distribution system contamination event. *Journal of Water Resources Planning and Management*, *143*(4), Article 04017001.

Swamee, P. K., & Sharma, A. K. (2008). *Design of water supply pipe networks*. John Wiley & Sons.

Thulasiraman, K., & Swamy, M. N. S. (2011). *Graphs: Theory and algorithms*. John Wiley & Sons.

UN-Water (2021). 2021: Summary progress update 2021–SDG 6–water and sanitation for all. URL https://www.unwater.org/sites/default/files/app/uploads/2021/12/SDG-6-Summary-Progress-Update-2021_Version-July-2021a.pdf.

Vrachimis, S. G., Eliades, D. G., Taormina, R., Kapelan, Z., Ostfeld, A., Liu, S., et al. (2022). Battle of the leakage detection and isolation methods. *Journal of Water Resources Planning and Management*, *148*(12), Article 04022068.

Vrachimis, S. G., Timotheou, S., Eliades, D. G., & Polycarpou, M. M. (2019). Iterative hydraulic interval state estimation for water distribution networks. *Journal of Water Resources Planning and Management*, *145*(1), Article 04018087.

Wächter, A., & Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, *106*(1), 25–57.

Yang, J., LeChevallier, M. W., Teunis, P. F., & Xu, M. (2011). Managing risks from virus intrusion into water distribution systems due to pressure transients. *Journal of Water and Health*, *9*(2), 291–305.