EKF-based observers for multi-leak diagnosis in branched pipeline systems

J.A. Delgado-Aguiñaga [a,*], I. Santos-Ruiz [b,d], G. Besançon [c], F.R. López-Estrada [b], V. Puig [d]

[a] Centro de Investigación, Innovación y Desarrollo Tecnológico CIIDETEC-UVM, Universidad del Valle de México, Campus Guadalajara Sur, CP 45601, Tlaquepaque, Jalisco, Mexico
[b] Tecnológico Nacional de México/Instituto Tecnológico de Tuxtla Gutiérrez, TURIX-Dynamics Diagnosis and Control Group, Carretera Panamericana km 1080 SN, Tuxtla Gutiérrez, México, CP 29050, Mexico
[c] Univ. Grenoble Alpes, CNRS, Grenoble INP, Institute of Engineering, GIPSA-lab, 38000 Grenoble, France
[d] Institut de Robòtica i Informàtica Industrial, CSIC-UPC, C/ Llorens i Artigas 4-6, 08028 Barcelona, Spain

ARTICLE INFO

Keywords:
Fault detection and isolation
Leak detection
Branching pipeline
Extended Kalman filter
Experimental results

ABSTRACT

The present work deals with the multi-leak diagnosis problem in a branched pipeline configuration as in water distribution systems. Here, it is assumed that the flow rate and pressure head measurements are available upstream and at all delivering points of the network. The proposed Leak Detection and Isolation (*LDI*) scheme basically involves two essential steps: leak *region* identification based on flow-rate residuals with a related k-Nearest Neighbors (k-*NN*) classifier, and then leak parameter identification (magnitude and position) via the use of the so-called Extended Kalman Filters (*EKFs*) for each leak based on a simple generic model and fed with pressure head estimations provided by an initial *EKF*. For the sake of illustration, successful experimental results of a two sequential leak scenario are provided using databases generated by a test bed plant with two branchings built at the Tuxtla Gutiérrez Institute of Technology.

1. Introduction

Pipeline systems are commonly used to transport fluids such as wastewater, drinkable water, and oil derivatives, among others, and they aim to satisfy human needs. A safe operation in the transporting process is always desired, however, abnormal behaviors such as leaks may occur and cause economic and environmental disasters. Leaks appear because of the natural aging of the pipe, earthquakes, external maintenance processes, or even illegal intrusion (gasoline pipelines).

Since the end of the last century, the leak diagnosis problem has been widely studied and several (*LDI*) strategies have been proposed for the single leak case in straight pipelines. In [1], a leak isolation methodology using a fitting loss coefficient calibration is proposed considering two stages: (a) the calculation of Equivalent Straight Length (*ESL*) by using an Extended Kalman Filter *EKF* as a state observer; (b) the implementation of an algebraic observer fed with the *ESL* estimated in the first stage. The estimation of the *ESL* provides the leak position in real coordinates, which is crucial in practice [2,3]. Experimental results demonstrate the effectiveness of this approach. Authors in [4] present a new *LDI* technique for pressurized pipelines based on an extension of a Differential Evolution (*DE*) algorithm. The leak localization problem is formulated as an optimization task using the classical dynamical model describing the fluid transient response inside a pipeline. The core of this approach is finding reasonable estimations related to the leak parameters while a cost function is minimized. Experimental findings of a pipeline prototype illustrated accurate results that were

* Corresponding author.
E-mail address: jorge.delgado@uvmnet.edu (J.A. Delgado-Aguiñaga).

then compared with estimations obtained from an *EKF*, which in turn, is widely used in this field. Following this direction, in [5] a parameter calibration process based on the *DE* algorithm is proposed and applied to the leak diagnosis problem. Experimental results show a good performance of the leak location task, including mapping of the calibration parameters that facilitate the searching process. Also, [6,7] describe how to estimate the roughness coefficient and the minor losses due to accessories, which is required to parameterize a model for leak diagnosis purposes either in single pipelines and pipeline networks. Based on heuristic methods, in [8] a bank of observers is proposed. Here, a pair of leak coefficients (magnitude and position) is taken from a search space and assigned to an observer. A Genetic Algorithm (*GA*) is exploited to minimize the integration of the square observation error such that the minimum integral observation error will be reached by the observer where the estimated leak parameters match the real values. Experimental results show a good performance of this algorithm. Finally, real-life problems have been successfully solved by applying leak diagnosis techniques as reported in [2]; here, authors highlight the arising difficulties in practice and they describe in detail the implementation of an *EKF* for locating a leak in an aqueduct situated in Guadalajara México.

The multi-leak problem has also been studied in the literature. In [9], both transient and static behavior of the fluid in leak condition are used to identify the parameters associated to the leaks without requirements of valve perturbation. The key of the method is the automatic selection of the specific family of models to be identified using the steady state conditions produced by the leaks. In [10], a multi-leak diagnosis strategy is proposed based on *EKFs*. The leaks are detected and identified as they appear. For the current leak, previous leak parameters are considered as constant variables. This experimental approach is applied successfully. Recently in [11], a scheme for detecting and locating multiple sequential leaks is proposed. This approach is based on using an adaptive observer to identify the hydraulic gradient in real-time and a leak location observer to estimate the leak position and its outflow. Experimental results of a pilot pipeline showed a good estimation despite operation changes and leaks.

On the other hand, in a water transportation network, the interconnection of pipelines results in complex configurations that can adopt the shape of branching pipeline systems. They can be used to carry water, possibly from natural water sources to treatment plants before being distributed to final consumers in cities or even in rural areas. In this paper, such branching pipeline configurations are considered the core of our proposal since they are quite common in water transportation networks. In some recent studies, this type of systems have been already studied. For example, in [12,13] a strategy based on a set of observers together with a logic detection function for isolating a leak in a pipeline is proposed. Here, it is assumed that the outflow in a branch junction is known from sensor measurements at the main pipeline ends. However, this method is limited to leaks occurring in the main pipe, while leaks occurring in the branch pipelines are not considered.

On the other hand, following this direction, frequency-based methods have also been reported. In [14], the leakage localization problem in tree-structured pipe networks using transient waves is studied on the basis of a frequency domain wave propagation model for a general tree-structured pipe network where the leak parameters (location and size) are factorized. The leak localization procedure is performed on the basis of the Matched-Field Processing (*MFP*) principle. This approach is evaluated via simulation and also by using a test-bed plant. Multiple leaks can be identified as long as the distance between them is larger than half the minimum wavelength of probing transient wave. More recently in [15], a frequency domain inverse transient analysis method for simultaneous identification of visco-elastic parameters and leaks in visco-elastic pipelines based on transfer matrix method is proposed. Such a method is applied to branched pipeline systems. Experimental results clearly demonstrate that visco-elasticity of the pipe wall is important for leak detection in visco-elastic pipes using the transient wave. Moreover, extensions to the multi-leak case have also been proposed; in [16] a procedure to detect multiple leaks for a specific element in pipe network systems through a detailed analysis of transient pressure signals based on an impedance matrix approach is presented. This approach considers a scenario of sequential leaks in which a decomposition scheme allows the multiple leak signal from the transient pressure response. However, the results presented are reported using a hypothetical heterogeneous pipe network system. It should be noted that in many real scenarios, the generation of pressure waves cannot be easily performed due to the lack of devices. Moreover, in case those pressure waves are somehow generated, they will produce water hammers that in turn could cause more leaks as discussed in [2], which is counterproductive.

Since the extension from a single pipeline to multiple pipes in a complex configuration brings new scientific challenges, the main objective of the present work is to address the multi-*LDI* problem in a branched pipeline system. This work can be seen as an extension of previous studies for single pipelines, mainly from our last work in [17], in which results via simulation were provided for the *single leak case*.

In this paper, the approach proposed combines *EKFs* with flow error residual analysis in a two-step method for multi-leak detection and identification in a branched pipeline. As a first step, the leak *region* is identified based on flow-rate residuals with a related *k-NN* classifier. Secondly, a first *EKF* provides estimations of pressure heads at all internal nodes, which are entered into a couple of *EKFs* in generic form that provide estimates of the leak parameters for each leak case: magnitude and position. This approach is fairly tractable in practice (because only two low-order *EKFs* are needed) and reliable because of some “unicity” of the family of leak signature patterns (guaranteed by an observability property of the system [18]). An important advantage of using a model-based approach as the proposed, is that any leak occurring throughout the system can be identified, and this methodology is not limited to specific locations, unlike other approaches based on steady-state simulations where leaks are typically assumed to occur in the nodes.

The proposed *LDI* approach is presented as a general leak diagnosis strategy for branching pipelines and for the sake of illustration, experimental results are presented by using a pilot plant system composed of a main pipe with two branches, which was built at the Tuxtla Gutiérrez Institute of Technology.

This work is organized as follows: Section 2 presents the modeling tools, including fluid dynamics in a pipeline, leak effect, and the related special boundary conditions needed for analysis in a branching configuration. The leak region identification on the basis of a *k-NN* classifier is presented in Section 3. Section 4 presents the leak diagnosis strategy and *EKF* designs. In Section 5, the experimental results with databases from the prototype of the Tuxtla Gutiérrez Institute of Technology are presented, and, finally, some conclusions and perspectives are given in Section 6.

2. Modeling of pipeline dynamics in a branched configuration

Let us consider a branching pipeline system just as shown in Fig. 1, where n_j with $j \in \{0, \ldots, \kappa+1\}$ stands for the j-th node and p_k with $k \in \{1, \ldots, \kappa+1\}$ stands for the k-th pipeline section linking two consecutive nodes in the *main pipeline*. Nodes n_0 and $n_{\kappa+1}$ stand for the boundary conditions at upstream and downstream in the main pipeline, respectively. B_m with $m \in \{1, \ldots, \kappa\}$ denotes the m-th branching which joins nodes N_m and n_m, accordingly.[1] In addition, in order to facilitate the overall dynamical modeling,

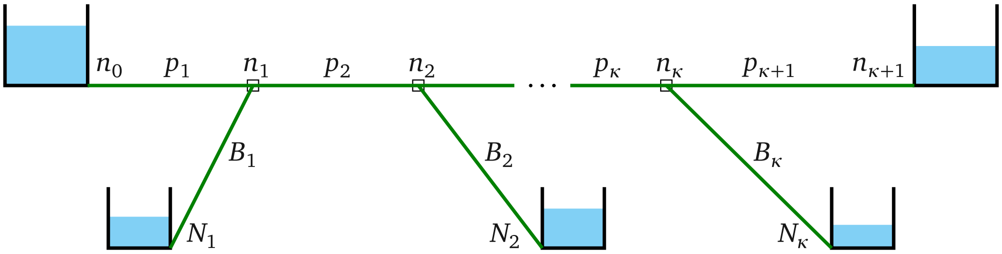


Fig. 1. Scheme of a κ-branched pipeline system.

each branching B_m will be redefined as a pipe section $p_{\kappa+1+m}$ in the modeling section:

$$B_m \longleftrightarrow p_{\kappa+1+m}. \tag{1}$$

For a κ branched pipeline system there are thus $2\kappa+1$ pipe sections.

The transient flow in a closed conduit is usually described by conservation mass and momentum equations referred to as Water Hammer Equations (*WHEs*), which are a pair of quasilinear hyperbolic Partial Differential Equations (*PDEs*). These *PDEs* are derived under the following assumptions: (i) the pipeline is considered to be straight, without fittings; (ii) the fluid is slightly compressible; (iii) the duct wall is slightly deformable; (iv) the convective velocity changes are negligible; (v) the cross-section and the fluid density are constant.

Thus, to analyze the dynamics in a branching configuration, let us firstly consider pipeline section p_κ linking nodes n_{k-1} and n_k; see Fig. 1. For this particular section, the nonlinear hyperbolic *PDEs* governing the fluid transient response can be written as [19]:

Momentum Equation

$$\frac{\partial Q(z_k,t)}{\partial t} + gA_k \frac{\partial H(z_k,t)}{\partial z_k} + \mu_k Q(z_k,t)\left|Q(z_k,t)\right| + g\ sin\ \alpha_k = 0 \tag{2}$$

Continuity Equation

$$\frac{\partial H(z_k,t)}{\partial t} + \frac{b_k^2}{gA_k}\frac{\partial Q(z_k,t)}{\partial z_k} = 0 \tag{3}$$

In these equations, distance z_k [m] and time t [s] are two independent variables, and pressure $H(z_k,t)$ [m] and flow rate $Q(z_k,t)$ [m^3/s] are two dependent variables. In addition, g is the gravity acceleration [m/s^2], and $\mu_k = f(Q_k)/(2D_kA_k)$ is the friction coefficient [20] depending on time and on parameters of the system f, b, A, and D, which may depend on space, α_k is the elevation angle of the section p_k. Here $z_k \in [0, L_k]$ denotes the position along the pipe p_k, and L_k is the equivalent straight length between nodes n_{k-1} and n_k, [21].

For this single pipeline analysis, boundary conditions for *PDEs* (2) and (3) are here considered to be

$$\begin{aligned} H(z_{k_0},t) &= H_{\text{in}}(t) \\ H(z_{k_{L_k}},t) &= H_{\text{out}}(t) \end{aligned} \tag{4}$$

for external functions $H_{\text{in}}(t)$ and $H_{\text{out}}(t)$, where z_{k_0}, $z_{k_{L_k}}$ refer to position 0 and L_k, respectively, along pipe section p_k. Hereinafter, the analysis being given for this section p_k, subscript k will be omitted for simplicity.

2.1. Leak model

A leak can appear arbitrarily at position $z_l \in (0, L)$, and it can be considered as a new boundary condition in *PDEs* (2), and (3) with outflow $Q_l(z_l,t) = C_d A_l \sqrt{2g}\sqrt{H_l(z_l,t)}$, in which C_d is the discharge coefficient and A_l is the leak cross-section area. Now by defining $\lambda \equiv C_d A_l \sqrt{2g}$, $Q_l(z_l,t)$ can be expressed as [22]

$$Q_l(z_l,t) = \lambda H_{t_l}(t)\sqrt{H_l(z_l,t)} \tag{5}$$

[1] It should be noticed that the proposed method cannot deal with the pipe systems with loop structures and the cases that more branches connected to $N_1, \ldots, N_k$.

in which $Q_l(z_l,t)$ is the leak flow rate [m^3/s], $H_l(z_l,t)$ is the pressure head [m] at the leak point, λ is the leak coefficient [m$^{5/2}$/s], and $\mathcal{H}_{t_l}(t)$ is the Heaviside step function associated with the leak occurrence at time t_l.

2.2. Spatial-discretization of the governing equations

On the other hand, to obtain a more tractable model for simulation and estimation purposes, a finite-dimensional description of *PDEs* (2) and (3) is considered. Here, the finite-difference method is used and the finite-dimensional approximation is as follows [19]:

$$\frac{\partial H(z_i,t)}{\partial z} \backsimeq \frac{H_{i+1}-H_i}{\Delta z_i} \quad \forall i=1,\ldots,n \tag{6}$$

$$\frac{\partial Q(z_{i-1},t)}{\partial z} \backsimeq \frac{Q_i-Q_{i-1}}{\Delta z_{i-1}} \quad \forall i=2,\ldots,n \tag{7}$$

Therefore, by using those finite difference approximations (6) and (7), the original *PDEs* (2) and (3) can be approximated by a pair of nonlinear ordinary differential equations keeping time as a continuous variable. Moreover, by considering boundary conditions (4) and also that a leak described by (5) may occur at the end of each pipe section, the pipeline length can be divided into n sections of sizes Δz_i $\forall i=1,\ldots,n$, with $\sum_{i=1}^{n}\Delta z_i = L$, to represent $n-1$ possible leaks. Thus, a finite-dimensional model for any number of sections can be obtained as follows [19] as in Fig. 2:

$$\dot{Q}_i = \frac{-gA}{\Delta z_i}\left(H_{i+1}-H_i\right) - \mu_i Q_i|Q_i| \quad \forall i=1,\ldots,n \tag{8}$$

$$\dot{H}_{i+1} = \frac{-b^2}{gA\Delta z_i}\left(Q_{i+1}-Q_i+\mathcal{H}_{t_{l_i}}\lambda_i\sqrt{H_{i+1}}\right) \quad \forall i=1,\ldots,n-1 \tag{9}$$

where H_1, H_{n+1} correspond to the boundary conditions H_{in}, H_{out} according to (4). Here $z_0=0$, $z_n=L$, and z_i for $i\neq(0,n)$ corresponds to an interior discretization node.

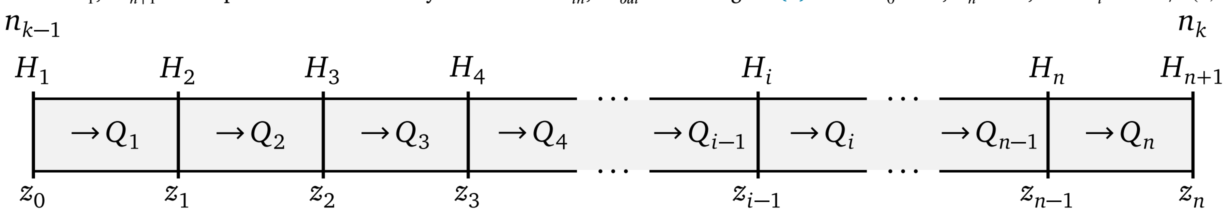


Fig. 2. Spatial discretization of the pipeline section p_k.

To extend the representation of the governing equations from a single pipeline system (8) and (9) to a more complex configuration like a branching pipeline system (made of interconnected single elements), special boundary conditions are in fact needed, which are described in the next subsection.

2.3. Special boundary conditions

2.3.1. Series junction

A series junction is given when two consecutive pipeline sections p_k and p_{k+1} are joined at node n_k, see Fig. 1. Such pipelines can have different diameters or represent a change of diameter in a pipeline at position z_i, see Fig. 3. This situation can also be considered when the wall thicknesses, wall materials, and/or friction factors are different.

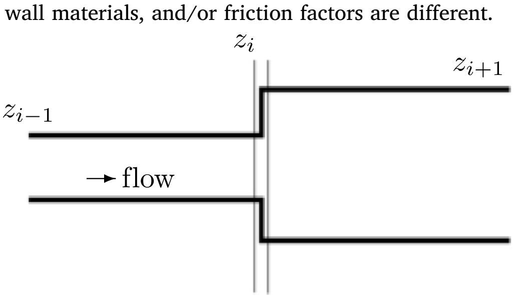


Fig. 3. Series junction in the pipeline section p_k at position z_i.

Note that if the head loss at the junction is neglected, then from the energy equation [19]: $H(z_{i^-},t)=H(z_{i^+},t)$. The continuity equation at both sides of the junction is $Q(\Delta z_i,t)=Q(\Delta z_{i+1},t)$.

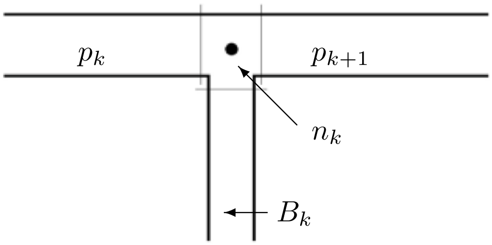


Fig. 4. Branched junction at node n_k.

2.3.2. Branch junction

For a branch junction at node n_k, see Fig. 4, the continuity equation is used, and it must be satisfied. In other words, there is no storage capacity at the branch junction. In addition, a common head is assumed when minor effects are neglected [23].

The flow balance at the node satisfies

$$Q_{p_k}(t) = Q_{p_{k+1}}(t) + Q_{B_k}(t) \tag{10}$$

with $H_{n_k}(t)$ a common pressure head at node n_k, as in [19,23], and with diameters possibly being different.

Remark 1. Since complex pipeline systems consist of series and branch connections, no new boundary condition is needed to handle complicated pipeline configurations [23].

2.4. Modeling for a branching pipeline system

Let us consider the case of a branching pipeline system with a constant head level upstream and at all delivery points (downstream), respectively. This scenario can represent a water transport system of primary elements, such as aqueducts that deliver a fluid to secondary locations like a potable water treatment plant; see Fig. 1. Corresponding boundary conditions are summarized in Table 1.

Let us also consider that a leak can occur between two consecutive nodes: n_k and n_{k+1} for a branching pipeline system with κ number of branches, just as shown in Fig. 1. For the main pipeline, a low order model can be obtained by connecting Eqs. (8) and (9) with $n = 2$ for each section p_k, and using Eqs. (5) for leaks or (10) for each branching accordingly, of the following form:

$$\dot{Q}_i = \frac{-gA}{\Delta z_i}\left(H_{i+1} - H_i\right) - \mu_i Q_i|Q_i| \quad \forall i = 1,\dots,2\kappa+2 \tag{11}$$

$$\dot{H}_{i+1} = \frac{-b^2}{gA\Delta z_i}\left(Q_{i+1} - Q_i + Q_{out_i}\right) \quad \forall i = 1,\dots,2\kappa+1 \tag{12}$$

here κ is the number of branchings, where

$$Q_{out_i} = \mathcal{H}_{t_i}\lambda_i\sqrt{H_{i+1}} \quad if\ i = 1,3,5,\dots,2\kappa+1 \tag{13}$$

or

$$Q_{out_i} = Q_{\eta+i-1} \quad if\ i = 2,4,6,\dots,2\kappa \tag{14}$$

with $\eta := 2\kappa + 2$. In the same way, for the mth branching, the following equations are obtained, for $m = 1,2,\dots,\kappa$:

$$\dot{Q}_{\eta+2m-1} = \frac{-gA}{\Delta z_{\eta+2m-1}}\left(H_{\eta+2m} - H_{2m+1}\right) - \mu_{\eta+2m-1}Q_{\eta+2m-1}|Q_{\eta+2m-1}| \tag{15}$$

$$\dot{H}_{\eta+2m} = \frac{-b^2}{gA\Delta z_{\eta+2m-1}}\left(Q_{\eta+2m} - Q_{\eta+2m-1} + Q_{out_{\eta+m-1}}\right) \tag{16}$$

$$\dot{Q}_{\eta+2m} = \frac{-gA}{\Delta z_{\eta+2m}}\left(H_{\eta+2m+1} - H_{\eta+2m}\right) - \mu_{\eta+2m}Q_{\eta+2m}|Q_{\eta+2m}| \tag{17}$$

with $Q_{out_{\eta+m-1}} = H_{t_{l_{2m}}}\lambda_{2m}\sqrt{H_{\eta+2m}}$. Fig. 5 shows the schematic diagram of such variable description that can be used for any κ branching pipeline system with a leak in each pipe section.

The model previously described can represent the occurrence of a leak in each pipeline section as they appear separately in time. The proposed leak diagnosis strategy considers two basic stages: leak region identification and leak parameter estimation. In the following subsection, the identification of the leaky pipe section is described in detail.

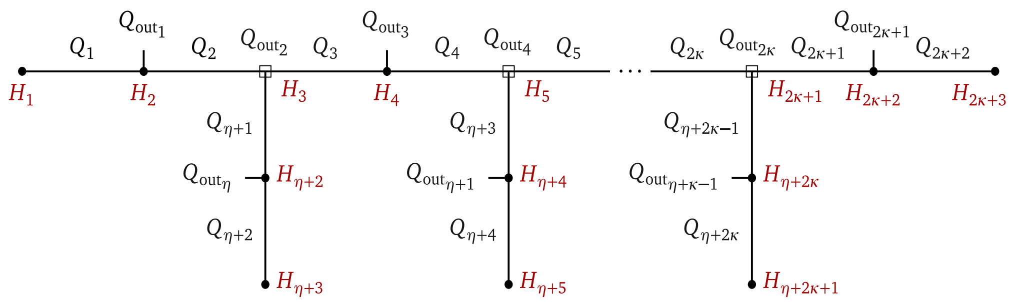


Fig. 5. Schematic diagram of a κ-branched pipeline system with leaks.

Table 1
Boundary conditions of the branching pipeline system.

Node	Variable	Availability
n_0	H_1	Via sensor
$n_1,\dots n_\kappa$	$H_3, H_5, \dots, H_{2\kappa+1}$	Immeasurable
	$H_2, H_4, \dots, H_{2\kappa+2}$	Immeasurable
$n_{\kappa+1}$	$H_{\eta+1}$	Via sensor
$N_1,\dots,N_\kappa$	$H_{\eta+3}, H_{\eta+5}, \dots, H_{\eta+2\kappa+1}$	Via sensor

3. Leak region identification by using a *k*-NN classifier

Solving the multi leak problem in a branching pipeline system can be an easy task if the region where each leak occurs is first identified: this makes the observer problem for leak parameter estimation of the lowest order. Thus, in a first step, we propose to identify the leak region based on directional residuals assuming that each new leak occurs in a different pipeline section. Such a region identification problem can be formulated as a multi-class classification problem, which can be solved using a machine learning technique like the k-Nearest Neighbors (k-*NN*).

To do that, one can consider a branching pipeline system as shown in Fig. 1, that is, with $\kappa+1$ sections in the main pipeline and κ branching. Each section/branching can be named as a class $S_1, S_2, \dots, S_{2\kappa+1}$. Thus, leaks occurring in the i-section are classified in the S_i class, and so on, see Fig. 6.

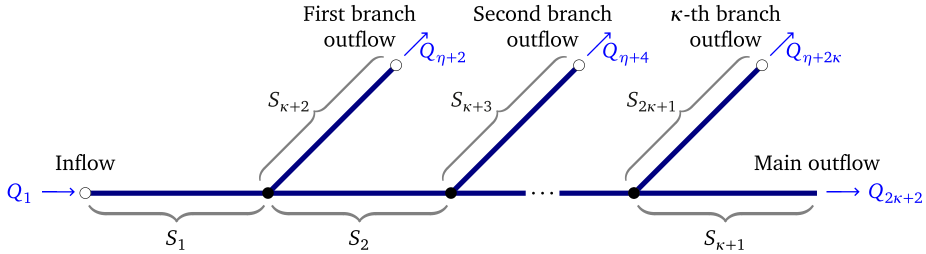


Fig. 6. Leak classification in a branched pipeline, with η as in Fig. 5.

On the other hand, flow rate measurements are available at upstream and downstream of the main pipeline $Q_{\text{mp}} = [Q_1\ Q_{2\kappa+2}] \in \mathbb{R}^2$ but also at all delivering points (downstream of each branching) $Q_{\text{br}} = [Q_{\eta+2}\ Q_{\eta+4}\ \dots\ Q_{\eta+2\kappa}] \in \mathbb{R}^\kappa$ and they are used as classification variables. The difference between flows from a faulty and a nominal system (leak-free) is known as residual. Those residuals are used as classifier inputs. A residual vector is built as:

$$\mathbf{r} = \mathbf{Q} - \mathbf{Q}^{(0)}, \tag{18}$$

where $\mathbf{Q} = [Q_{\text{mp}}\ Q_{\text{br}}]^T \in \mathbb{R}^{\kappa+2}$ represents the flow rate measurements (with leakage) and $\mathbf{Q}^{(0)}$ represents the nominal flow rates (without leakage).

For different leaks of the same class S_κ and under their magnitude variations, the characteristic direction in the $(\kappa+2)$-dimensional space does not deviate and thus leak classes S_i can be distinguished.

In practice, the measured variables are corrupted by some noise. In the case of the S_i class, the residuals' directions are not the same, but they tend to cluster within a roughly conical shaped region. The region's exact geometry is difficult to determine analytically because it depends, among other things, on the noise level and the pipeline topology (the number of branching and the physical parameters). This makes challenging the identification of the section where each leak occurs only by comparing the residual vector direction with predefined limits. For this reason, the use of a k-*NN* classifier is proposed that considers supervised learning and correctly classifies all leaks according to their corresponding class by viewing the directional proximity between leaks in the same class. Moreover, if a new leak occurs while a previous one is active (with parameters identified), the *k-NN* classifier still works well.

In the proposed *k-NN* classifier, the similarity between residual vectors is not measured with the classical definition of Euclidean distance, but the so-called *cosine distance*. This method depends on the direction of the vectors rather than their magnitudes. For any two vectors, $\mathbf{r}$ and $\mathbf{r}'$, the cosine distance is defined by:

$$d_{\cos}(\mathbf{r},\mathbf{r}') = 1 - \frac{\mathbf{r}\cdot\mathbf{r}'}{\|\mathbf{r}\|\,\|\mathbf{r}'\|} = 1 - \frac{\sum_{k=1}^{\kappa+2} r_k r'_k}{\sqrt{\sum_{k=1}^{\kappa+2} r_k^2}\sqrt{\sum_{k=1}^{\kappa+2} (r'_k)^2}} \tag{19}$$

The *k-NN* classifier's objective is to find among different groups of known leaks which class is more consistent with the direction of a given new residual (new leak), assuming that a smaller cosine distance between residuals indicates a considerable similarity. The classification has two stages: training and prediction.

1. The *training* of the *k-NN* classifier is an offline process. In this process, a set of residual samples corresponding to leaks of available classes and their labels $(S_1, S_2, \dots, S_{2\kappa+1})$ is stored. The dataset to train the classifier is obtained by simulating leaks at different positions in each pipeline section S_i. For example, the training dataset can be built by simulating leaks at 10 evenly distributed positions in each section, giving a total of $10\times(2\kappa+1)$ different leaks.
2. Leak class *prediction* is an online process. Here, a continuous comparison of the most recent residual with the labeled residuals from the training dataset is performed. If the leak class is denoted by S, and $\Pr\left(S = S_i|\mathbf{r}\right)$ is the probability that the leak corresponds to the S_i class given the residual $\mathbf{r}$, the k-NN classifier assumes that

$$\Pr\left(S = S_i|\mathbf{r}\right) = \frac{k_i}{k}, \tag{20}$$

where k_i is the number of residuals in the ith class among the k nearest neighbors to the residual $\mathbf{r}$. The cosine distance expressed in (19) is used as a metric to quantify the nearness between residuals. The class with the highest probability is chosen as the output of the classifier:

$$\hat{S} = \arg\max_i \Pr\left(S = S_i|\mathbf{r}\right) \tag{21}$$

The effectiveness of the *k-NN* classifier is evaluated by using a test data set by calculating the fraction of correctly classified leaks. The number of nearest neighbors to use, k, is determined by cross-validation, maximizing the classifier's accuracy on the test data.

Remark 2. Several distance metrics were evaluated: cosine, Euclidean, Chebyshev, Manhattan, Minkowski and Mahalanobis. However the performance of the k-NN classifier using the cosine distance as a metric of similarity showed a significant improvement in comparison with the others. In particular, the Euclidean distance produced in average 13% misclassifications whereas poor results were obtained by using the other distance metrics: Chebyshev, Manhattan, Minkowski and Mahalanobis.

4. Leak diagnosis strategy

4.1. Leak detection

Since we are trying to solve the non-concurrent leak problem in a branching pipeline system, it is assumed that several leaks appear sufficiently separate in time. It is considered that each new leak occurs in a different branching (pipe section). The outflow of the first leak can be estimated as follows, see Fig. 6:

$$Q_{leak_1}(t) = \left(Q_1(t) - \left(Q_{2\kappa+2}(t) + \sum_{i=1}^{\kappa} Q_{\eta+2i}(t)\right)\right) \tag{22}$$

when the right-hand side is larger than some threshold δ_1 chosen to indicate the leak presence considering the noise variance of measurements, its value should be chosen to avoid false alarms. If $Q_{leak_1}(t) < \delta_1$ it is considered that there is no leak (δ_1 is approximated by inspection as explained later on). The flow of the subsequent leaks can be computed as

$$Q_{leak_j}(t) = \left(Q_1(t) - \left(Q_{2\kappa+2}(t) + \sum_{i=1}^{\kappa} Q_{\eta+2i}(t) + \sum_{i=1}^{j-1} \tilde{Q}_i\right)\right) \quad j = 2, \dots, p \tag{23}$$

where $\tilde{Q}_i$ is the mean outflow of each previous leak and computed in a time window between the time occurrence of leak j and leak $j+1$. With this strategy, it is possible to know the flow of each leak, which helps to reduce the complexity of observer design for leak parameter identification.

4.2. Leak parameter identification

In order to estimate the parameters of each leak, we use the information previously obtained in the leak region and leak flow rate $Q_{leak_j}(t)$.

Firstly, unmeasured pressure heads at the inner nodes of main pipeline, that is, at nodes $n_1, \ldots, n_\kappa$ (see Table 1), are estimated. This is done on the basis of model (11)–(12), with known inputs H_1, $H_{2\kappa+3}$, measured outputs Q_1, $Q_{2\kappa+2}$ and estimated flow rates Q_{out_i}, from Q_{leak_j}, Eqs. (13)–(14) and Eq. (16) in a steady state.

Notice that the case when leaks only appear in branchings is more specific to the configuration under study here, and for this reason, let us present it in more detail: in that case, $Q_{2i} = Q_{2i-1}$ for $i = 1$ to $\kappa + 1$ in main pipeline, and model (11)–(12) reduces to equations for Q_{2i+1}, $i = 0$ to κ, and H_{2i+1}, $i = 0$ to κ, that is:

$$\begin{bmatrix} \dot{Q}_1 \\ \dot{H}_3 \\ \vdots \\ \dot{Q}_{2i+1} \\ \dot{H}_{2i+3} \\ \vdots \\ \dot{Q}_{2\kappa+1} \end{bmatrix} = \begin{bmatrix} \frac{-gA}{\Delta z_1+\Delta z_2}\left(H_3 - H_1\right) - \mu_1 Q_1|Q_1| \\ \frac{-b^2}{gA(\Delta z_1+\Delta z_2)}\left(Q_3 - Q_1 + Q_{out_2}\right) \\ \vdots \\ \frac{-gA}{\Delta z_{2i+1}+\Delta z_{2i+2}}\left(H_{2i+3} - H_{2i+1}\right) - \mu_{2i+1} Q_{2i+1}|Q_{2i+1}| \\ \frac{-b^2}{gA(\Delta z_{2i+1}+\Delta z_{2i+2})}\left(Q_{2i+3} - Q_{2i+1} + Q_{out_{2i}}\right) \\ \vdots \\ \frac{-gA}{\Delta z_{2\kappa+1}+\Delta z_{2\kappa+2}}\left(H_{2\kappa+3} - H_{2\kappa+1}\right) - \mu_{2\kappa+1} Q_{2\kappa+1}|Q_{2\kappa+1}| \end{bmatrix}, \tag{24}$$

where the positions of branchings along the main pipeline are considered to be known a priori, and $[H_1\ H_{2\kappa+3}]^T$ and $[Q_1\ Q_{2\kappa+1}]^T$ are the measured input and output, respectively. Such a model can be used for estimation of pressure heads H_{2i+1} (see next section).

Then, we can move to estimate the leak parameters: assuming for instance the occurrence of a leak event in branching m, i.e. in region $S_{\kappa+1+m}$, see Fig. 6, we can now refer to model (15)–(17). For the purpose of estimating leak position $\Delta z_{\eta+2m-1}$ and leak magnitude λ_{2m}, this model can be rewritten in an extended form as:

$$\begin{bmatrix} \dot{Q}_{\eta+2m-1} \\ \dot{H}_{\eta+2m} \\ \dot{Q}_{\eta+2m} \\ \dot{\Delta z}_{\eta+2m-1} \\ \dot{\lambda}_{2m} \end{bmatrix} = \begin{bmatrix} \frac{-gA}{\Delta z_{\eta+2m-1}}\left(H_{\eta+2m} - H_{2m+1}\right) - \mu_{\eta+2m-1} Q_{\eta+2m-1}|Q_{\eta+2m-1}| \\ \frac{-b^2}{gA\Delta z_{\eta+2m-1}}\left(Q_{\eta+2m} - Q_{\eta+2m-1} + \lambda_{2m}\sqrt{H_{\eta+2m}}\right) \\ \frac{-gA}{L_m-\Delta z_{\eta+2m-1}}\left(H_{\eta+2m+1} - H_{\eta+2m}\right) - \mu_{\eta+2m} Q_{\eta+2m}|Q_{\eta+2m}| \\ 0 \\ 0 \end{bmatrix} \tag{25}$$

where L_m is the length of branching m, $u = [H_{2m+1}\ H_{\eta+2m+1}]$ is used as input, and $y = [Q_{\eta+2m-1}\ Q_{\eta+2m}]$ is used as output. In practice, H_{2m+1} is replaced by an estimate $\hat{H}_{2m+1}$ obtained in the previous estimation step, and $Q_{\eta+2m-1}$ is obtained by steady state estimation:

$$\begin{bmatrix} Q_1 \\ Q_3 \\ Q_5 \\ \vdots \\ Q_{2i+1} \\ \vdots \\ Q_{2\kappa-1} \\ Q_{2\kappa+1} \end{bmatrix} = \begin{bmatrix} Q_{in} \\ Q_1 - Q_{out_2} \\ Q_3 - Q_{out_4} \\ \vdots \\ Q_{2i-1} - Q_{out_{2i}} \\ \vdots \\ Q_{2\kappa-3} - Q_{out_{2\kappa-2}} \\ Q_{2\kappa-1} - Q_{out_{2\kappa}} \end{bmatrix}. \tag{26}$$

Both structures (24) and (25) with considered measurements can be represented in a compact form as:

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= Hx \end{aligned} \tag{27}$$

for some nonlinear functions f, g and matrix H, with u corresponding to driving pressure heads, and y measured flows. Notice that both models (24) and (25) satisfy an observability condition so that one can expect possible state reconstruction from available measurements (see Appendix A).

4.2.1. Kalman filter design

Let us present here an actual solution for state estimation of the models (24), and (25) based on the Kalman filter [24]. Since the implementation in practice implies using sampled variables provided by the data acquisition system, here a discrete-time Kalman filter as a state observer is used. Both systems, (24) and (25), under form (27) are discretized in time, by using Heun's method because of its good trade-off between simplicity and accuracy [10]. The discrete-time version of (27) can be rewritten in the following form:

$$\begin{aligned} x^{k+1} &= \xi_d(x^k, u^k, u^{k+1}) \\ y^k &= Hx^k \end{aligned} \tag{28}$$

where k is the index of discrete time.

A first *EKF* is then designed for system (24) expressed in the form (28). A second *EKF* is designed for system (25) also expressed in the form (28). For both discretized systems, state x, output y and input u are summarized in Table 2:

Table 2
State variables: inputs and outputs for Kalman designs.

	State x	Output y	Input u
EKF 1	$x = [x_1\ x_2\ \cdots\ x_{2\kappa}\ x_{2\kappa+1}]^T =$ $[Q_1\ H_3,\ \cdots\ H_{2\kappa+1}\ Q_{2\kappa+1}]^T$	$y = [y_1\ y_2]^T =$ $[Q_1\ Q_{2\kappa+1}]^T$	$u = [u_1\ u_2]^T =$ $[H_1\ H_{2\kappa+3}]^T$
EKF 2	$x = [x_1\ x_2\ x_3\ x_4\ x_5]^T =$ $[Q_{\eta+2m-1}\ H_{\eta+2m}\ Q_{\eta+2m}\ \Delta z_{\eta+2m-1}\ \lambda_{2m}]^T$	$y = [y_1\ y_2]^T =$ $\left[\hat{Q}_{\eta+2m-1}\ Q_{\eta+2m}\right]^T$	$u = [u_1\ u_2]^T =$ $\left[\hat{H}_{2m+1}\ H_{\eta+2m+1}\right]^T$

Table 3
Kalman filter equations.

$\hat{x}^{k^-}$ is the a priori estimate of x^k: $\hat{x}^{k^-} = \xi_d(\hat{x}^{k-1}, u^{k-1})$	P^{k^-} is the a priori covariance matrix: $P^{k^-} = F^k P^{k-1}(F^k)^T + Q$
K^k is the Kalman gain: $K^k = P^{k^-}H^T(HP^{k^-}H^T + \mathcal{R})^{-1}$	P^k is the a posteriori covariance matrix: $P^k = (I - K^k H)P^{k^-}$
F^k is the Jacobian matrix: $F^k = \frac{\partial \xi_d(x,u)}{\partial x}\|_{x=\hat{x}^k}$	$\mathcal{R}$ and Q are the covariance matrices of measurement and process noises $\mathcal{R} = \mathcal{R}^T > 0$, and $Q = Q^T > 0$ $P^{0^-} = (P^{0^-})^T > 0$

Table 4
Prototype pipeline sensors-devices: *FT* and *PT* stand for *Flow Transducer* and *Pressure Transducer*, respectively.

Sensor/Device	Trademark	Specifications
(FT), *Promag Proline* 10 *P*	*Endress Hauser*™	4 – 20 [mA]
(PT), *PMP* 41	*Endress Hauser*™	4 – 20 [mA]
$(PT_{L1}$ and $PT_{L3})$	*Winters*™	4 – 20 [mA]
Pump 1, 7502*MEAU*	*Siemens*™	5 [HP]
Pump 2, 3*HME*100	*Evans*™	1 [HP]

Table 5
Pipeline parameters.

Parameter	Symbol	Value
Pipeline diameter	D	0.0486 [m]
Pipeline length main pipeline	L	84.58 [m]
Pipeline length branching 1	L_{B_1}	30.6 [m]
Pipeline length branching 2	L_{B_2}	30.6 [m]
Relative roughness	ε	3.47×10^{-4}
Fluid kinematic viscosity	ν	8.03×10^{-7} [m^2/s]
Fluid density	ρ	996.59 [kg/m^3]
Acceleration due to gravity	g	9.79 [m/s^2]
Pressure wave speed	b	422.75 [m/s]

The state of *EKF* is given by [24]

$$\hat{x}^k = \hat{x}^{k^-} + K^k(y^k - H\hat{x}^{k^-}), \tag{29}$$

where K is the correction gain. Full equations are given in Table 3.

The implementation of both kind of Kalman observers are fully described in Appendix B.

5. Experimental results

5.1. Experimental setup

The effectiveness of the algorithm previously described was verified experimentally in a pipeline prototype built at the hydraulics laboratory of the National Technological Institute of Mexico (TecNM) in Tuxtla Gutiérrez, see Fig. 7. The system parameters are shown in Table 5. The pipeline prototype is manufactured with a 2-inch inner diameter PVC pipe. Through it, drinkable water is transported under pressure driven by a 5 [HP] centrifugal pump, although a variable frequency device can regulate the driving power. The pipeline is designed to operate as a looped pipeline, but an open/close valve configuration can modify it to adopt the structure of the main pipeline with two branches (as in our case study). The main pipe is assembled in a serpentine shape at a lower level, while the branches are also serpentine-shaped but at an upper-level, 1 [m] above the main pipeline.

In the databases, flow rate and pressure head measurements are saved from sensors (sensor-device information can be seen in Table 4) placed at the ends of the system, through the data acquisition card DAQ modular Yokogawa GM10 and at a sampling rate of 100 [Hz]. Finally, the discrete-time extended Kalman filter is tested offline in a *MATLAB*™ environment.

The simplified Process and Instrumentation diagram (P&I) of the system, which is configured to operate in a branched pipeline mode, is shown in Fig. 8. Pressure and flow rate sensors/transmitters are available upstream and at all delivery points. Also, pressure sensors are also available *near* the branch junctions (namely, PT05 and PT06). However, it should be noted that these measurements are not used as inputs for the proposed algorithm, but they are used to validate the results. To compensate the elevation effect, 1 [m] of pressure head is added to all values of the PT04 and PT03 measurements.

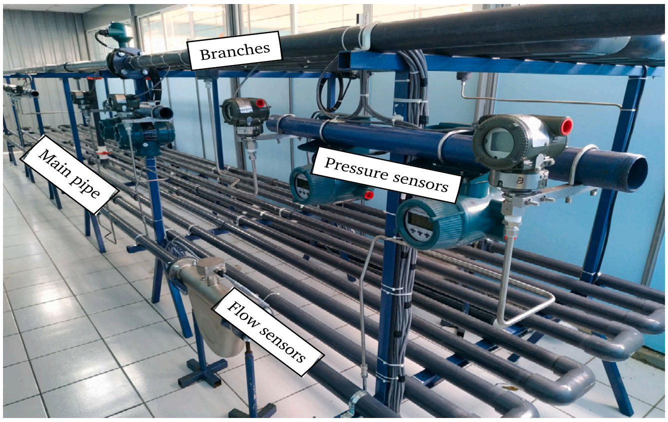


Fig. 7. Prototype pipeline at TecNM in Tuxtla Gutiérrez, Mexico.

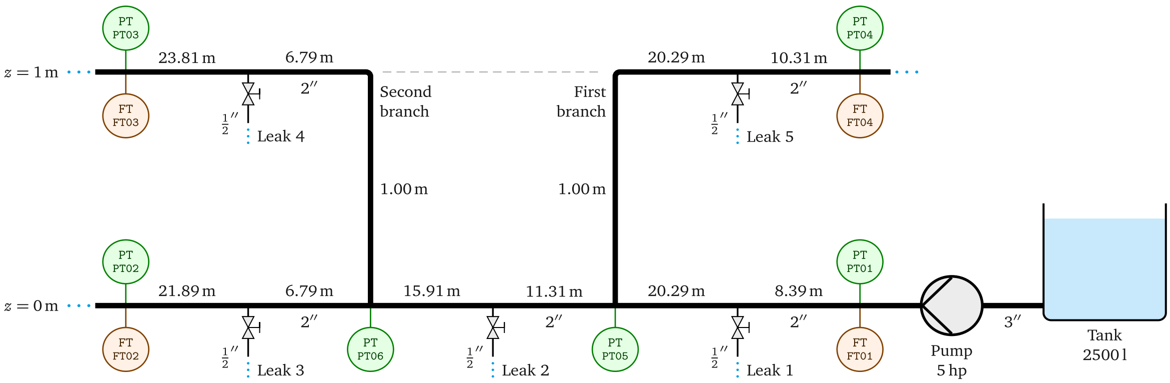


Fig. 8. P&I diagram of the prototype pipeline.

The valves labeled *Leak 1*, *Leak 2*, and *Leak 3* can be used to cause leaks in the main pipeline, whereas the valves labeled *Leak 4* and *Leak 5* can be used to cause leaks in each branching, respectively. With $\kappa = 2$, a dynamical description can be obtained using Eqs. (11) up to (17), and the distribution of variables is shown in Fig. 9:

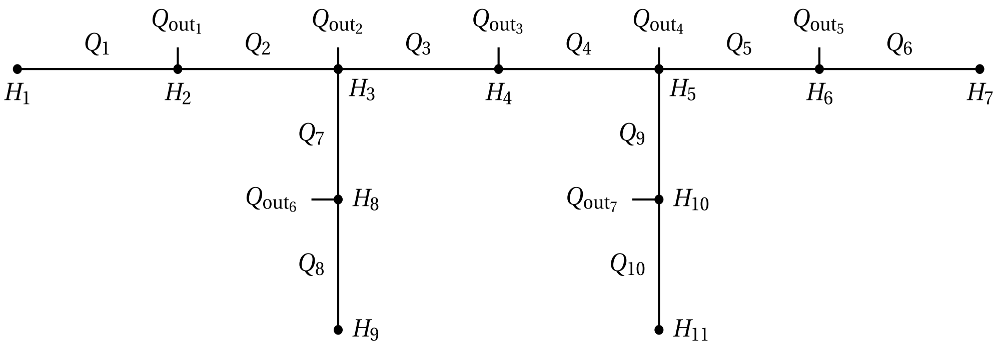


Fig. 9. Variable description of a two-branched pipeline system.

Notice that Q_{out_1}, Q_{out_3} and Q_{out_5} will be different from zero if valves *Leak* 1, *Leak* 2 and *Leak* 3 are opened, but since we focus here on leaks in branchings, they will be closed.

For an experimental validation of the proposed method, a case of two non-concurrent leaks (corresponding to valves *Leak* 4 and *Leak* 5, Q_{out_6} and Q_{out_7}, respectively) is considered and evaluated in the test bed previously described, considering two different situations:

- (i) the first leak occurs in branching 2 (*Leak 5*: Q_{out_7}) and the second in branching 1 (*Leak 4*: Q_{out_6});

• (ii) the first leak occurs in branching 1 (*Leak 4*: Q_{out_6}) and the second one in branching 2 (*Leak 5*: Q_{out_7}).

To do that, a database is created with the data acquisition card DAQ modular Yokogawa GM10 at a sampling rate of 100 [Hz], and sensors installed in the pilot plant. The experiment design is as follows:

1. The experiment starts in a nominal condition (without leak). At time $t = t_1 \gg 0$, the valve labeled *Leak 5* (Q_{out_7}) is opened.
2. At time $t_2 \gg t_1$, a second leak is caused by opening valve *Leak 4* (Q_{out_6}). Both leaks are maintained until the end of the experiment.
3. The second experiment is performed in the same way as the first one but the valve *Leak 4* is firstly opened.
4. A database is generated and saved in leak-free condition under the same sampling rate and duration as in the faulty ones.

5.2. Leak detection and region identification

The occurrence of each leak can be easily detected using (22) and (23). Figs. 12 and 13 show the total outflow of each experiment. However, from these graphics, it is not possible to identify the region where each leak is occurring. Thus, the identification of the sequence of leak events is first performed on the basis of the *k-NN* classifier described in Section 3. For the first experiment, the sequence of events is as follows: a leak in class S_5 at time t_1 and a leak in class S_4 at time t_2, see Fig. 10.

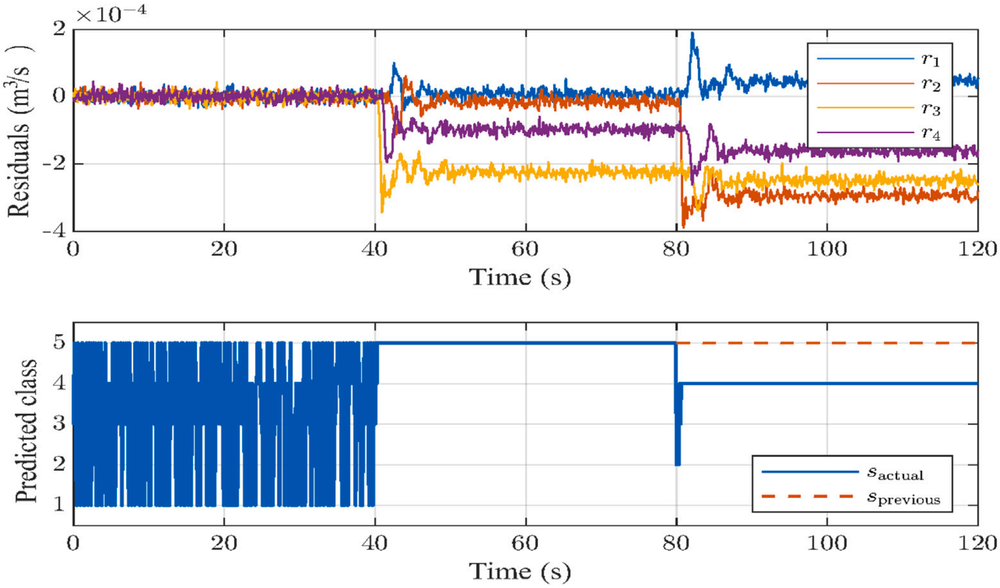


Fig. 10. Identification of the leak class S_5 at time t_1 and S_4 at time t_2, $t_2 > t_1$.

For the second experiment, the sequence of events is as follows: a leak in class S_4 at time t_1 and a leak in class S_5 at time t_2, see Fig. 11.

Now, by knowing the sequence of events from graphics shown in Figs. 12 and 13, the values of threshold δ and the time of the leaky event t_i can be approximated: (a) From the first experiment (see Fig. 12) $t_1 \approx 120$ [s], $t_2 \approx 240$ [s], $\delta_1 = 1.41 \times 10^{-5}$ [m^3/s] and $\delta_2 = 4.19 \times 10^{-4}$ [m^3/s]. (b) From the second experiment (see Fig. 13) $t_1 \approx 82$ [s], $t_2 \approx 170$ [s], $\delta_1 = 1.4 \times 10^{-5}$ [m^3/s] and $\delta_2 = 3.986 \times 10^{-4}$ [m^3/s].

Since the values of $\delta's$ and the time of the leaky events have been identified, the flow of each leak can be easily computed by means of Eqs. (22) and (23), Figs. 14 and 15 depict such a leak flow separation.

Remark 3. The leak region identification (pre-localization) stage is performed in steady state. In our case study, the k-NN classifier accurately determined the *region* of the network where the leak is occurring almost instantaneously, since the duration of the flow transients is short due to the small length of the pipes. However, for real distribution networks (larger and more complex pipeline systems), the flow transients will be longer and consequently, there will be a delay in the leak region identification stage, i.e., until a new steady state is reached. On the other hand, to obtain enough data of all possible leak scenarios to correctly train the k-NN classifier, a well-calibrated hydraulic model is required. However, in practice, obtaining such a calibrated hydraulic model is not easily achieved.

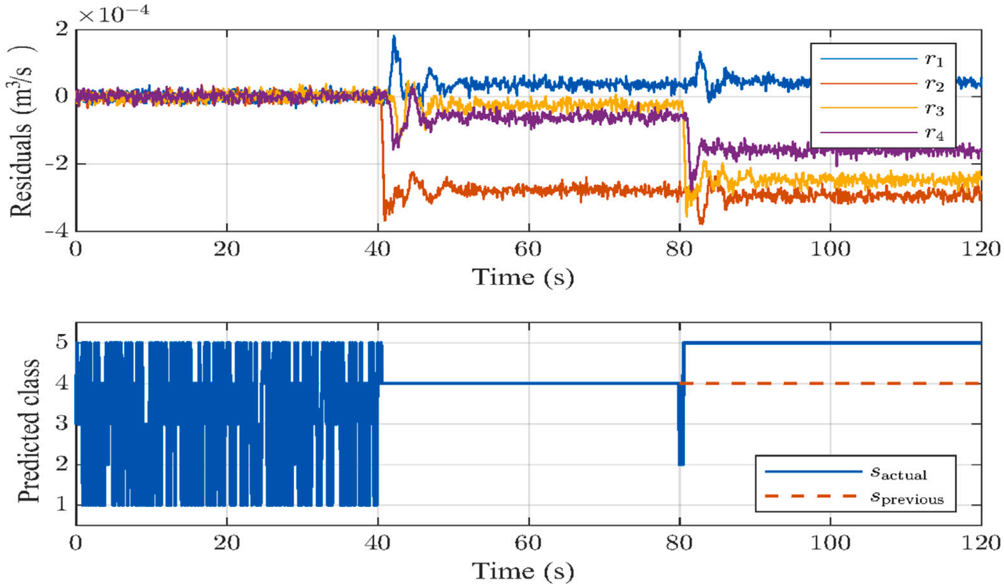


Fig. 11. Identification of the leak class S_4 at time t_1 and S_5 at time t_2, $t_2 > t_1$.

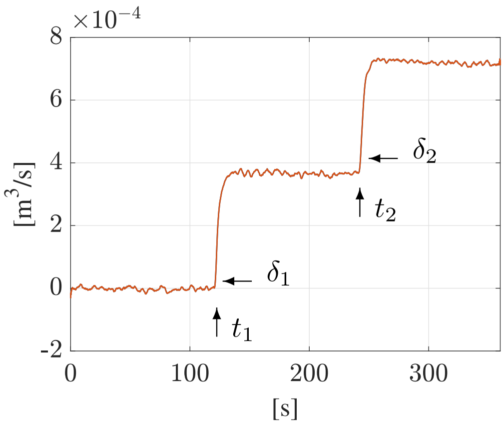


Fig. 12. Total outflow of sequential leaky events of experiment 1: S_5 (*valve: Leak 5*) at t_1 and S_4 (*valve: Leak 4*) at t_2, $t_2 > t_1$.

5.3. Leak parameter identification

5.3.1. Estimation of pressure head and flow rate in the main pipeline

By using the steady state relations shown in (26), the following equations are obtained:

$$\begin{bmatrix} Q_1 \\ Q_3 \\ Q_5 \end{bmatrix} = \begin{bmatrix} Q_{in} \\ Q_1 - Q_{out_2} \\ Q_3 - Q_{out_4} \end{bmatrix}. \tag{30}$$

where Q_{out_2} and Q_{out_4} are obtained as follows: (a) for the first experiment:

$$\begin{aligned} Q_{out_2} &= Q_7 = Q_8 + Q_{leak_2} \\ Q_{out_4} &= Q_9 = Q_{10} + Q_{leak_1} \end{aligned} \tag{31}$$

(b) for the second experiment:

$$\begin{aligned} Q_{out_2} &= Q_7 = Q_8 + Q_{leak_1} \\ Q_{out_4} &= Q_9 = Q_{10} + Q_{leak_2} \end{aligned} \tag{32}$$

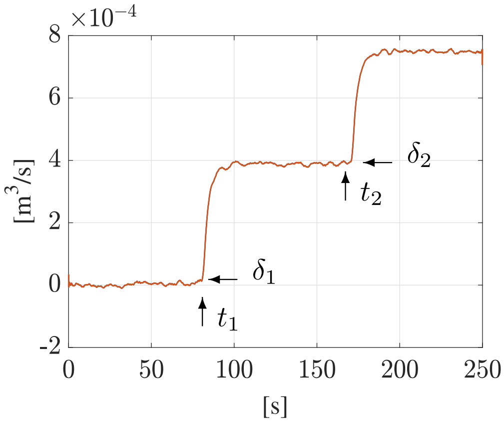


Fig. 13. Total outflow of sequential leaky events of experiment 2: S_4 (*valve: Leak 4*) at t_1 and S_5 (*valve: Leak 5*) at t_2, $t_2 > t_1$.

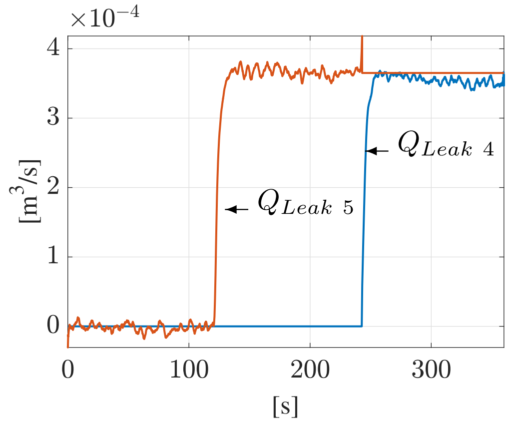


Fig. 14. Separation of total outflow, experiment 1.

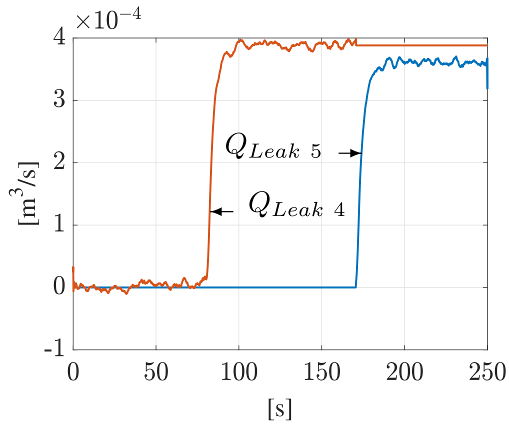


Fig. 15. Separation of total outflow, experiment 2.

Notice that Q_{leak_1} and Q_{leak_2} were identified in the previous procedure. To estimate the pressure heads H_3 and H_5 (interior nodes), for each experiment an *EKF* is designed for system (33) which is derived directly from (24)

$$\begin{bmatrix} \dot{Q}_1 \\ \dot{H}_3 \\ \dot{Q}_3 \\ \dot{H}_5 \\ \dot{Q}_5 \end{bmatrix} = \begin{bmatrix} \frac{-gA}{\Delta z_1+\Delta z_2}\left(H_3 - H_1\right) - \mu_1 Q_1|Q_1| \\ \frac{-b^2}{gA(\Delta z_1+\Delta z_2)}\left(Q_3 - Q_1 + Q_{out_2}\right) \\ \frac{-gA}{\Delta z_3+\Delta z_4}\left(H_5 - H_3\right) - \mu_3 Q_3|Q_3| \\ \frac{-b^2}{gA(\Delta z_3+\Delta z_4)}\left(Q_5 - Q_3 + Q_{out_4}\right) \\ \frac{-gA}{\Delta z_5+\Delta z_6}\left(H_7 - H_5\right) - \mu_5 Q_5|Q_5| \end{bmatrix}, \tag{33}$$

Table 6
EKF initialization, main pipeline, experiment 1 and 2.

Variable	Value/case 1	Value/case 2	Units
$\hat{Q}_1^0$	0.0042	0.0042	[m^3/s]
$\hat{H}_3^0$	2.98	2.95	[m]
$\hat{Q}_3^0$	0.0022	0.0032	[m^3/s]
$\hat{H}_5^0$	1.86	1.83	[m]
$\hat{Q}_5^0$	0.0011	0.0012	[m^3/s]

where Q_{out_2} and Q_{out_4} are obtained by using (31) (resp. (32)). For both experiments, the initialization of the state is given in Table 6.

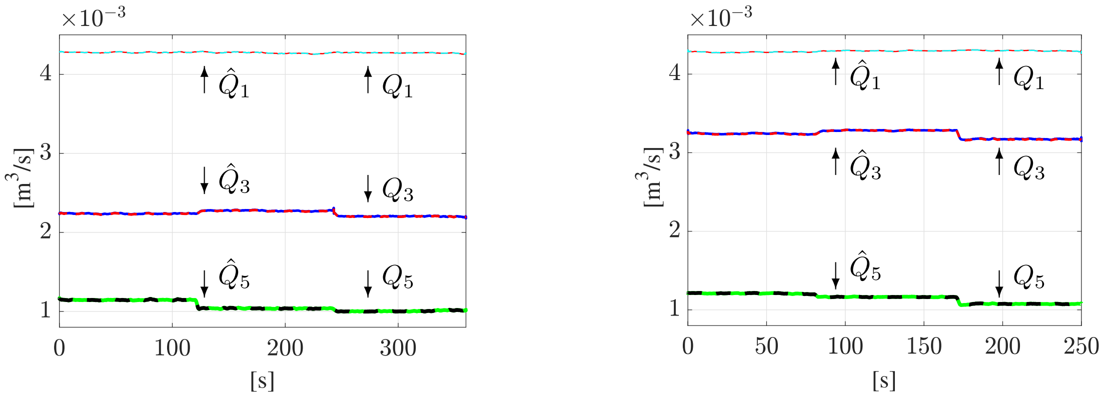
The covariance matrices of measurement and process noises for the *EKF* used for both experiments are respectively chosen as

$$\mathcal{Q} = \mathrm{diag}\left([1\times10^{-8}, 1\times10^{-1}, 1\times10^{-8}, 1\times10^{-1}, 1\times10^{-8}]\right) \tag{34}$$

$$\mathcal{R} = \mathrm{diag}\left([1\times10^{-8}, 1\times10^{-8}]\right). \tag{35}$$

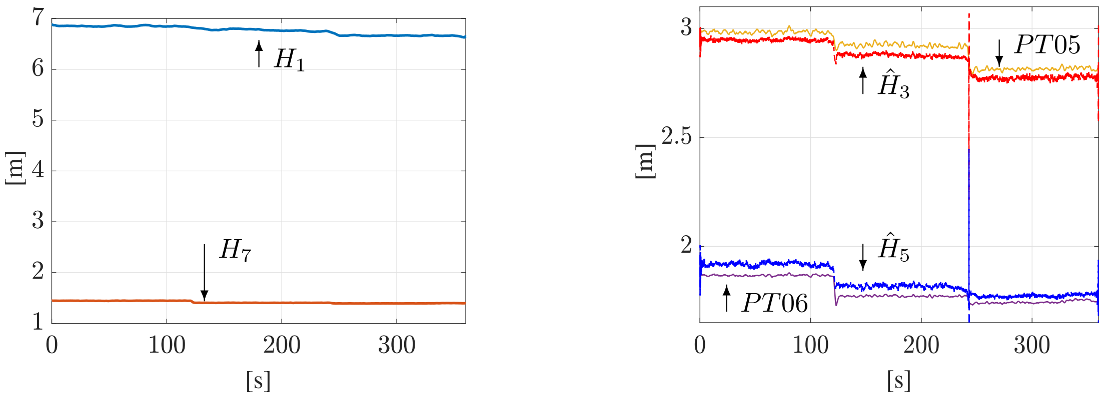
Hereinafter, the results obtained from the two experiments are presented. In Fig. 16(a) (resp. 16(b)), the estimation of the flow rate in the main pipeline (for system (33)) is depicted showing a good fitting. In the same way, in Figs. 17(a) and 17(b) (resp. 18(a) and 18(b)), a comparison between the estimated pressure head and those coming from measurements is shown. In Fig. 17(b) (resp. 18(b)), a detailed view is presented, and it can be observed that the estimations are not equal to the measurements because the sensors are not exactly in the branch position (where they are estimated). This means that the observer estimates the pressure head correctly as expected.

At this point, it is possible to identify the leak parameters since all boundary conditions in the branchings are known, i.e., $[H_9 \;\; H_{11}]$ via sensor measurement and $[\hat{H}_3 \;\; \hat{H}_5]$ via estimation.



(a) Estimation of flow rate in the main pipeline, experiment 1. (b) Estimation of flow rate in the main pipeline, experiment 2.

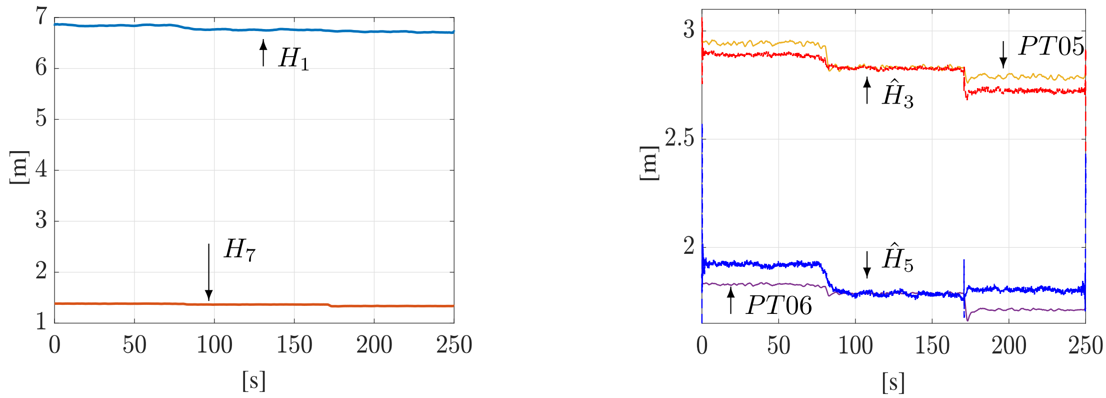
Fig. 16. Flow rate in the main pipeline.



(a) Pressure head at the ends of the main pipeline, experiment 1. (b) Estimation of pressure head at inner nodes, experiment 1.

Fig. 17. Pressure head in the main pipeline, experiment 1.

With those estimations given by the first *EKF* observer, it is possible to identify the parameters of both leaks for each experiment.



(a) Pressure head at the ends of the main pipeline, experiment 2 (b) Estimation of pressure head at inner nodes, experiment 2.

Fig. 18. Pressure head in the main pipeline, experiment 2.

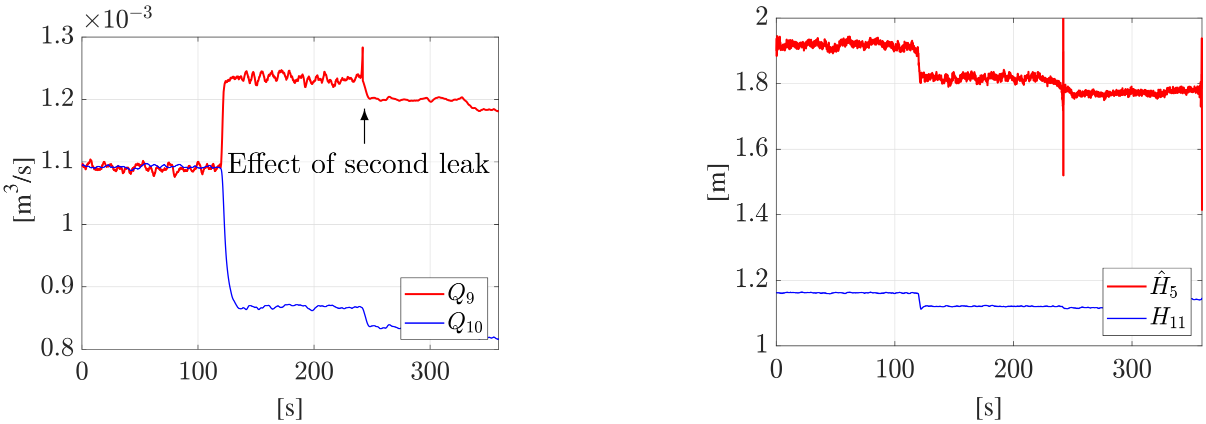
Table 7
EKF initialization, second branch.

Variable	Value	Units
$\hat{Q}_9^0$	1.08×10^{-3}	[m^3/s]
$\hat{H}_{10}^0$	1.31	[m]
$\hat{Q}_{10}^0$	1.08×10^{-3}	[m^3/s]
$\hat{\Delta z}_{p_2}^0$	19.12	[m]
$\hat{\lambda}_2^0$	0	[m$^{5/2}$/s]

5.4. Leak parameter estimation: experiment 1

5.4.1. Identification of parameters of leak 1 in second branching

The boundary conditions at the ends of p_5 are depicted in Fig. 19(b), and the flow rates are shown in Fig. 19(a). Furthermore, Q_9 is estimated as in (31) and $\hat{H}_5$ is provided via estimation by the *EKF* 1 (for system (24)).



(a) Estimated flow rate in the second branch. (b) Pressure head in the second branch.

Fig. 19. Flow rate and pressure head in the second branch, experiment 1.

A second and generic *EKF* is then applied for identifying the leak parameters (with the same structure as system (25) but with Δz_{p_2} and magnitude λ_2, respectively). The initialization state is shown in Table 7.

Similarly, the covariance matrices of measurement and process noises for the generic *EKF* are chosen as

$$\mathcal{Q} = \mathrm{diag}\left([1 \times 10^{-5},\ 1 \times 10^{-3},\ 1 \times 10^{-5},\ 1 \times 10^{6},\ 1 \times 10^{-7}]\right) \tag{36}$$

$$\mathcal{R} = \mathrm{diag}\left([1 \times 10^{-5},\ 1 \times 10^{-5}]\right). \tag{37}$$

Table 8
EKF initialization, first branch.

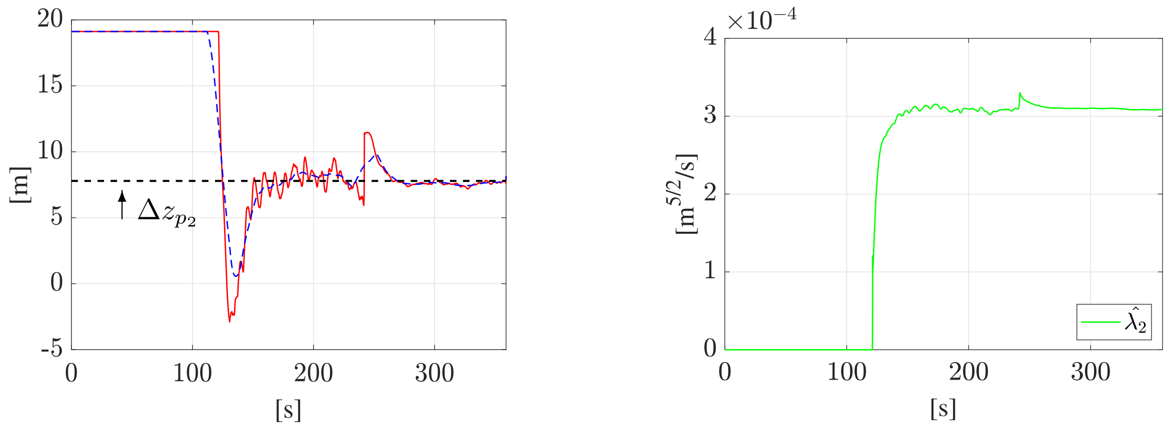
Variable	Value	Units
$\hat{Q}_7^0$	1.99×10^{-3}	[m^3/s]
$\hat{H}_8^0$	2.58	[m]
$\hat{Q}_8^0$	1.99×10^{-3}	[m^3/s]
$\hat{\Delta z}_{p_1}^0$	6.32	[m]
$\hat{\lambda}_1^0$	0	[m$^{5/2}$/s]

To obtain moving average values of the leak position, the estimation generated by the *EKF* was filtered with the equation:

$$\Delta_{z_F}(k) = \frac{1}{2N+1}(\Delta_{z_F}(k+N) + \Delta_{z_F}(k+N-1) + \cdots + \Delta_{z_F}(k-N)) \tag{38}$$

where $\Delta_{z_F}(k)$ is the smoothed value for the variable at time k, N is the number of neighboring data taken on either side of $\Delta_{z_F}(k)$, and $2N+1$ is the span dimension. A span equal to 101 was used for the position parameter.

The leak parameters: position and magnitude are depicted in Figs. 20(a) and 20(b), respectively.



(a) Estimation of leak position: Filtered version $\hat{\Delta} z_{p_{2_F}}$ dotted blue signal and original version $\hat{\Delta} z_{p_2}$ solid red signal.

(b) Estimation of leak coefficient.

Fig. 20. Leak parameters in the second branch, experiment 1.

5.4.2. Parameter identification of leak 2 in first branching

A third and generic *EKF* (similar to the previous case) is applied for identifying the leak parameters of the second leak. Here Q_7 is computed as in (32), and $\hat{H}_3$ is provided by the *EKF* 1. This third *EKF* is initialized using the states presented in Table 8, where Δz_{p_1} and λ_1 refer to the leak position and magnitude, respectively.

Similarly, the covariance matrices of measurement and process noises for this generic *EKF* are chosen as

$$\mathcal{Q} = \text{diag}\left([1 \times 10^{-5}, 1 \times 10^{-3}, 1 \times 10^{-5}, 5 \times 10^{3}, 1 \times 10^{-8}]\right) \tag{39}$$

and $\mathcal{R}$ equal to (37). The pressure head and flow rate at the ends of the first branch are shown in Fig. 21(b), and 21(a), respectively. The parameters of the leak are depicted in Figs. 22(a) and 22(b). In the same way as before, accurate results can be observed.

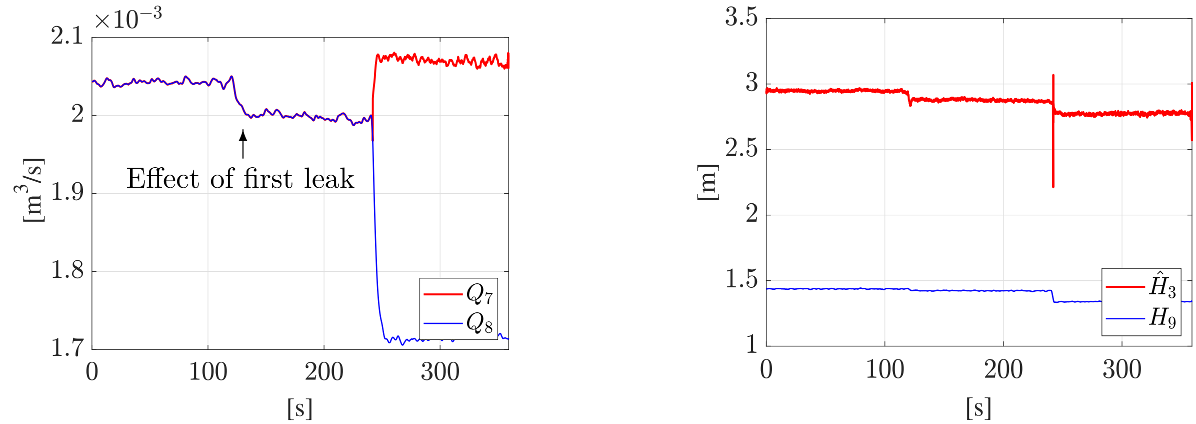
5.5. Leak parameter estimation: experiment 2

5.5.1. Parameter identification of leak 1 in first branching

In the same way as before, the boundary conditions at the ends of p_4 are depicted in Fig. 23(b), and the flow rates are shown in Fig. 23(a). Furthermore, Q_7 is estimated as in (32) and $\hat{H}_3$ is also provided via estimation by the *EKF* 1.

The second *EKF* is then again used for identifying the leak parameters on the basis of system (25) but with Δz_{p_1} and λ_1, respectively. The initialization state is shown in Table 9.

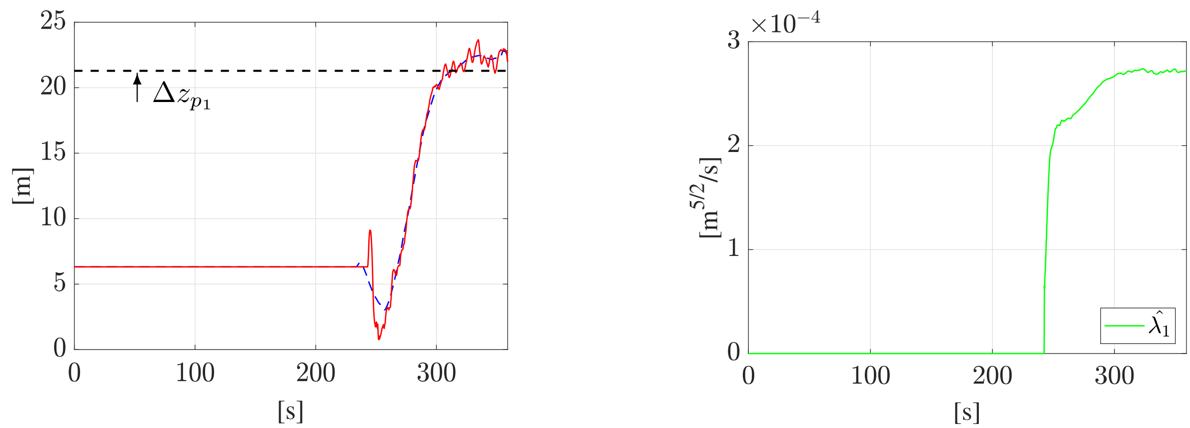
The covariance matrices of measurement and process noises are the same as (36) and (37). The parameters of the leak are depicted in Figs. 24(a) and 24(b).



(a) Flow rate.

(b) Pressure head.

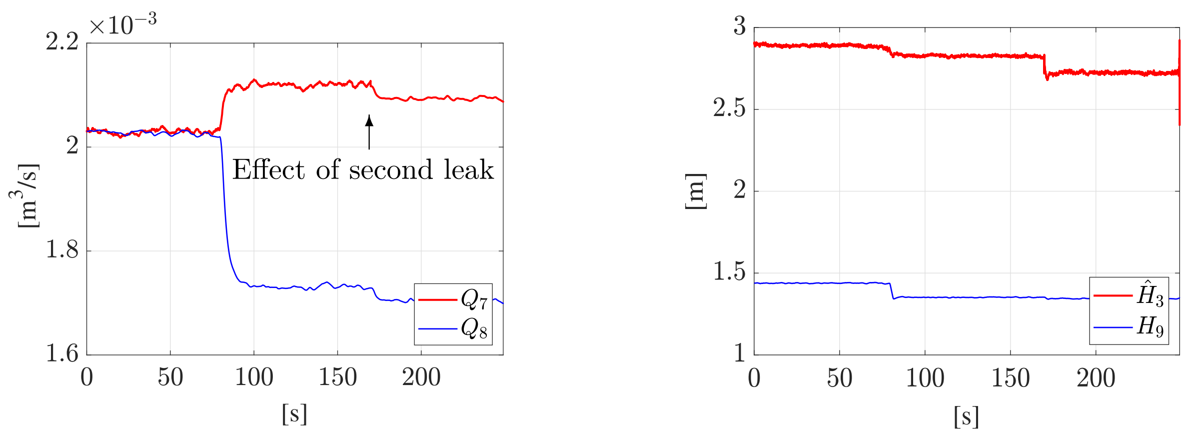
Fig. 21. Flow rate and pressure head in the first branch, experiment 1.



(a) Estimation of leak position: Filtered version $\hat{\Delta} z_{p_{1_F}}$ dotted blue signal and original version $\hat{\Delta} z_{p_1}$ solid red signal.

(b) Estimation of leak coefficient.

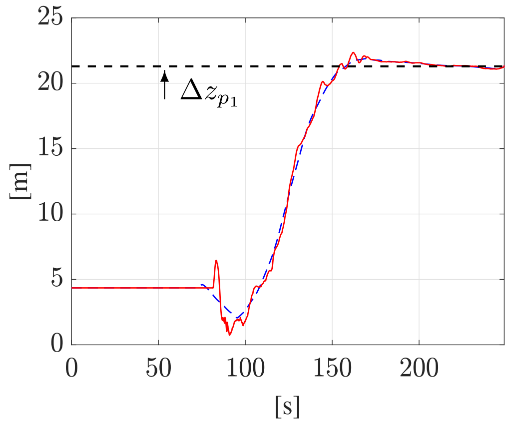
Fig. 22. Leak parameters in the first branch, experiment 1.



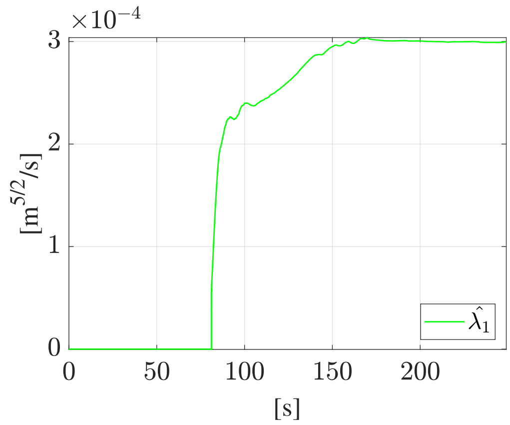
(a) Estimated flow rate in the first branch.

(b) Pressure head in the first branch.

Fig. 23. Flow rate and pressure head in the first branch, experiment 2.



(a) Estimation of leak position: Filtered version $\hat{\Delta} z_{p_{1_F}}$ dotted blue signal and original version $\hat{\Delta} z_{p_1}$ solid red signal.



(b) Estimation of leak coefficient.

Fig. 24. Leak parameters in the first branch, experiment 2.

Table 9
EKF initialization, first branch.

Variable	Value	Units
$\hat{Q}_7^0$	2.11×10^{-3}	[m^3/s]
$\hat{H}_8^0$	2.6	[m]
$\hat{Q}_8^0$	2.11×10^{-3}	[m^3/s]
$\hat{\Delta z}_{p_1}^0$	4.35	[m]
$\hat{\lambda}_1^0$	0	[m$^{5/2}$/s]

Table 10
EKF initialization, second branch.

Variable	Value	Units
$\hat{Q}_9^0$	1.13×10^{-3}	[m^3/s]
$\hat{H}_{10}^0$	1.07	[m]
$\hat{Q}_{10}^0$	1.13×10^{-3}	[m^3/s]
$\hat{\Delta z}_{p_2}^0$	17.25	[m]
$\hat{\lambda}_2^0$	0	[m$^{5/2}$/s]

5.5.2. *Parameter identification of leak 2 in second branching*

By following the same observer design as before, the *EKF* is initialized here using the states presented in Table 10.

Similarly, the covariance matrices of measurement and process noises for this generic *EKF* are chosen as (37) and (39), respectively. The pressure head and flow rate at the ends of the second branch are shown in Fig. 25(b) and Fig. 25(a), respectively. The parameters of the leak are depicted in Figs. 26(a) and 26(b). Once again, accurate results can be observed.

5.6. *Discussion of the results*

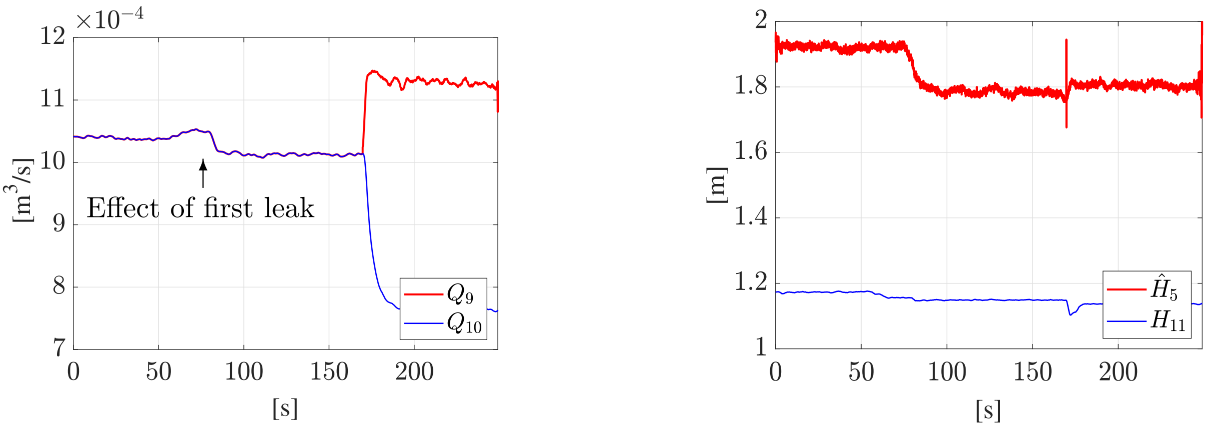
5.6.1. *Experiment 1*

Figs. 20(a), 22(a) show the leak position estimation for the first leak appearing in branching 2 and for the second leak appearing in branching 1, respectively. In Table 11, a quantitative index is presented to show the performance of both cases on the basis of the error norm criteria, which is computed as follows:

$$\|e_z\| = \sqrt{\sum_{j=1}^{Y-1} (e_z(j))^2} \tag{40}$$

where $e_z = \Delta z - \hat{\Delta z}$ and Y is the length of the vector.

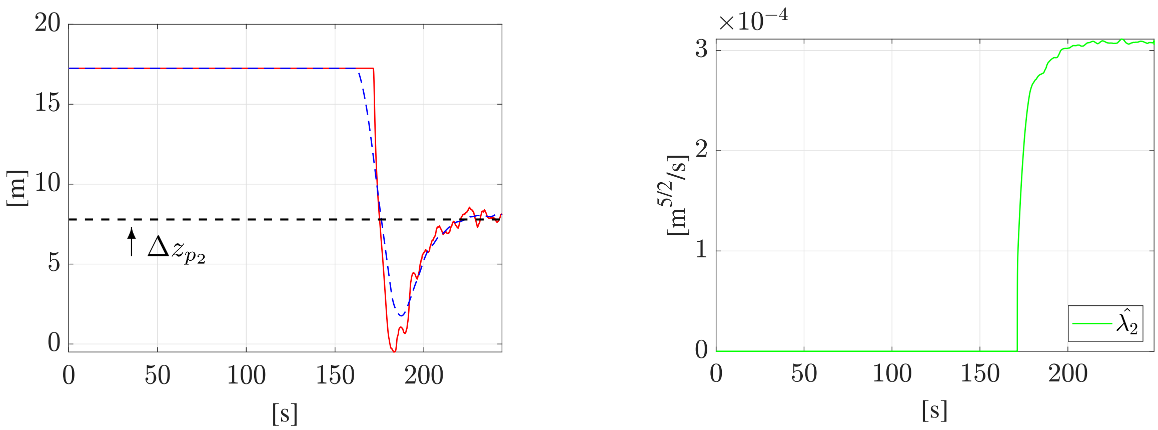
On the other hand, estimations of the leak magnitude are also depicted in Figs. 20(b) and 22(b), respectively. Several experiments were performed and in Table 12, the computational time spent by each *EKF* is depicted in statistical terms of mean value and standard deviation.



(a) Flow rate.

(b) Pressure head.

Fig. 25. Flow rate and pressure head in the first branch, experiment 2.



(a) Estimation of leak position: Filtered version $\hat{\Delta}z_{p2_F}$ dotted blue signal and original version $\hat{\Delta}z_{p2}$ solid red signal.

(b) Estimation of leak coefficient.

Fig. 26. Leak parameters in the second branch, experiment 2.

Table 11
Error norm for each case, experiment 1.

Case	Branching	$\|e_z\|$ [m]
First leak	2	0.0721×10^2
Second leak	1	0.1873×10^2

Table 12
Performance of the *EKFs* in computational-execution terms, experiment 1.

EKFs	Mean	Standard deviation
EKF 1, main pipeline	5.01 [s]	0.455 [s]
EKF 2, first leak	4.20 [s]	0.168 [s]
EKF 2, second leak	3.62 [s]	0.147 [s]

5.6.2. Experiment 2

In the same way, Figs. 24(a), 26(a) show the leak position estimation for the first leak appearing in branching 1 and for the second leak appearing in branching 2, respectively. In the same way as before, in Table 13, a quantitative index is presented to show the performance of both cases on the basis of the error norm criteria.

Table 13
Error norm for each case, experiment 2.

Case	Branching	$\|e_z\|$ [m]
First leak	1	0.0545×10^2
Second leak	2	0.1645×10^2

Table 14
Performance of the *EKFs* in computational-execution terms, experiment 2.

EKFs	Mean	Standard deviation
EKF 1, main pipeline	4.95 [s]	0.376 [s]
EKF 2, first leak	4.35 [s]	0.153 [s]
EKF 2, second leak	3.23 [s]	0.162 [s]

Similarly, estimations of the leak magnitude are also depicted in Figs. 24(b) and 26(b), respectively. In Table 14, the computational time spent by each *EKF* is summarized.

6. Conclusions

A methodology to detect and locate sequential leaks in a branched pipeline system has been proposed with successful experimental results. The approach is based on observer techniques (extended Kalman filters) but in a two-step procedure that significantly reduces the size of the observer problem. First, an error pattern analysis is used to identify the leak region, and then only two reduced-order observers are designed. The methodology is quite general and can address more complex branched pipeline configurations than the one considered in the paper, remaining easy to implement in practice. Its reliability has been experimentally validated on various scenarios, and further validation with real applications will be part of future developments.

Acknowledgments

This work was achieved during a postdoctoral research stay of the first author at Grenoble Images Parole Signal Automatique Laboratory of the Université Grenoble Alpes, *GIPSA-Lab*, Grenoble, France. The first author would like to thank deeply Prof. Gildas Besançon for give him the opportunity to develop this research stay in *GIPSA-Lab*.
We acknowledgment CONACyT for its support via the postdoctoral fellowship 244269 (for the first author) and the National Problems grant PN-2016/3595, but also *CIIDETEC-UVM*, *IRI-UPC*, Tecnológico Nacional de México *TNM* project: 14515.22-P and *TURIX*-Dynamics Diagnosis and Control Group Tuxtla Gutiérrez for their fruitful collaboration in this research.

Appendix A. Observability analysis

Let us consider systems (24) and (25) separately.

For system (24), a necessary condition for designing an observer is that such a system satisfies an appropriate *observability condition* [18], which can be checked to hold here independently of the inputs. Assuming that each Q_{out_i} is known, it can be checked that system (A.1) made of state variables x, up to $x_{2\kappa+1}$ ($\kappa+1$ is the number of pipe sections) satisfies the structure of uniform observability [18,25].

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \vdots \\ \dot{x}_{2\kappa} \\ \dot{x}_{2\kappa+1} \end{bmatrix} = \begin{bmatrix} \frac{-gA}{\Delta z_1}\left(x_2 - u_1\right) - \mu_1 x_1 |x_1| \\ \frac{-b^2}{gA\Delta z_1}\left(x_3 - x_1 + Q_{out_1}\right) \\ \frac{-gA}{\Delta z_2}\left(x_4 - x_2\right) - \mu_2 x_3 |x_3| \\ \frac{-b^2}{gA\Delta z_2}\left(x_5 - x_3 + Q_{out_2}\right) \\ \vdots \\ \frac{-b^2}{gA\Delta z_{2\kappa}}\left(x_{2\kappa+1} - x_{2\kappa-1} + Q_{out_{2\kappa}}\right) \\ \frac{-gA}{\Delta z_{2\kappa+1}}\left(u_2 - x_{2\kappa+1}\right) - \mu_{2\kappa+1} x_{2\kappa+1} |x_{2\kappa+1}| \end{bmatrix}. \tag{A.1}$$

System (A.1) with measurement x_1 can be rewritten under form

$$\begin{aligned} \dot{x} &= f_0(x) + u_1 f_1(x) + u_2 f_2(x) \\ y &= x_1 \end{aligned} \tag{A.2}$$

with state $x(t) \in \mathbb{R}^{2\kappa+1}$, input $u(t) = [u_1\ u_2]^T = [H_{in}(t)\ H_{out}(t)]^T \in \mathbb{R}^2$, and output $y(t) = x_1 = Q_{in}(t) \in \mathbb{R}$, for appropriate functions $f_i's$. Such a system in turn can be transformed into the following canonical uniformly observable structure

$$\begin{aligned} \dot{z} &= Az + \psi_0(z) + \psi_1(z)u_1 + \psi_2(z)u_2 \\ y &= Cz \end{aligned} \tag{A.3}$$

$$A = \begin{bmatrix} 0 & 1 & & 0 \\ \vdots & & \ddots & \\ & & & 1 \\ 0 & & \dots & 0 \end{bmatrix}, \psi_0(z) = \begin{bmatrix} 0 \\ \vdots \\ \psi_{2\kappa-1}(z) \end{bmatrix},$$

$$\psi_k(z) = \begin{bmatrix} \psi_{k_1}(z_1) \\ \vdots \\ \psi_{k_j}(z_1, \dots, z_j) \\ \vdots \\ \psi_{k_{2\kappa-1}}(z) \end{bmatrix}, C = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}^T \tag{A.4}$$

by means a diffeomorphism given by:

$$\Phi(x) = \begin{bmatrix} \Phi_1(x) \\ \Phi_2(x) \\ \vdots \\ \Phi_{2\kappa-1}(x) \end{bmatrix}. \tag{A.5}$$

where:

$$\Phi_k(x) = L_{f_0}^{k-1}(h)(x) \ for \ k = 1, \dots, 2\kappa - 1. \tag{A.6}$$

Similarly, for system (25) it was shown in [26] that observability is also satisfied in this case. By combining both properties, the overall system is observable.

Remark 4. The observability property can be guaranteed as long as the state estimation between two consecutive leaks is solved fast enough so that the leak is identified before the second leak appears [10].

Appendix B. Extended Kalman Filter implementation

Algorithm 1: Extended Kalman Filter 1

Input: $Q_{leak_1}, \dots, Q_{leak_p}$; H_1, $H_{2\kappa+3}$; Q_1 and $Q_{2\kappa+1}$.

Output: $\hat{x} = [\hat{x}_1 \ \hat{x}_2, \ \dots, \ \hat{x}_{2\kappa+1}]^T = [\hat{Q}_1 \ \hat{H}_3 \ \hat{Q}_3, \ \dots, \ \hat{H}_{2\kappa+1} \ \hat{Q}_{2\kappa+1}]^T$.

initialization: $\hat{x}^0 = E(x_0)$, $P^0 = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]$

k is the discrete time index; $Ts = 100$ [Hz] sampling time and T_f is the experiment duration.

while $kT_s <= T_f$ **do**

a) Compute the following partial derivative matrices

$$F^k = \frac{\partial \xi_d(x, u)}{\partial x}|_{x=\hat{x}^k} \tag{B.1}$$

b) Perform the time update of the state estimate and estimation-error covariance as follows

$$P^{k^-} = F^k P^{k-1} (F^k)^T + Q \tag{B.2}$$

$$\hat{x}^{k^-} = \xi_d(\hat{x}^{k-1}, u^{k-1}) \tag{B.3}$$

c) Compute the following partial derivatives

$$H^k = \frac{\partial H_d}{\partial x}|_{x=\hat{x}^{k^-}} \tag{B.4}$$

d) Perform the measurement update of the state estimate and estimation error covariance as follows

$$\begin{aligned} K^k &= P^{k^-} H^T (H P^{k^-} H^T + \mathcal{R})^{-1} \\ \hat{x}^k &= \hat{x}^{k^-} + K^k [(y^k - H \hat{x}^{k^-})] \\ P^k &= (I - K^k H) P^{k^-} \end{aligned} \tag{B.5}$$

$k = k + 1$

end

Algorithm 2: Extended Kalman Filter 2

Input: $y = [\hat{Q}_{\eta+2m-1}\ Q_{\eta+2m}]^T$; $u = [\hat{H}_{2m+1}\ H_{\eta+2m+1}]^T$; δ_i.
Output: $\hat{x} = [\hat{x}_1\ \hat{x}_2\ \hat{x}_3\ \hat{x}_4\ \hat{x}_5]^T = [\hat{Q}_{\eta+2m-1}\ \hat{H}_{\eta+2m}\ \hat{Q}_{\eta+2m}\ \Delta z_{\eta+2m-1}\ \lambda_{2m}]^T$.
initialization: $\hat{x}^0 = E(x_0)$, $P^0 = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]$
k is the discrete time index; $Ts = 100$ [Hz] sampling time and T_f is the experiment duration.
while $kT_s <= T_f$ **do**
  if $Q_{leak_i} > \delta_i$ **then**
    a) Compute the following partial derivative matrices

$$F^k = \frac{\partial \xi_d(x,u)}{\partial x}|_{x=\hat{x}^k} \tag{B.7}$$

    b) Perform the time update of the state estimate and estimation-error covariance as follows

$$P^{k^-} = F^k P^{k-1} (F^k)^T + \mathcal{Q} \tag{B.8}$$

$$\hat{x}^{k^-} = \xi_d(\hat{x}^{k-1}, u^{k-1}) \tag{B.9}$$

    c) Compute the following partial derivatives

$$H^k = \frac{\partial H_d}{\partial x}|_{x=\hat{x}^{k^-}} \tag{B.10}$$

    d) Perform the measurement update of the state estimate and estimation error covariance as follows

$$\begin{aligned} K^k &= P^{k^-} H^T (H P^{k^-} H^T + \mathcal{R})^{-1} \\ \hat{x}^k &= \hat{x}^{k^-} + K^k[(y^k - H\hat{x}^{k^-})] \\ P^k &= (I - K^k H) P^{k^-} \end{aligned} \tag{B.11}$$

    $k = k + 1$
  else
    $\hat{x} = \hat{x}_0$;
  end
end

References

[1] A. Navarro, O. Begovich, J. Sánchez, G. Besançon, Real-time leak isolation based on state estimation with fitting loss coefficient calibration in a plastic pipeline, Asian Journal of Control 19 (1) (2017) 255–265, http://dx.doi.org/10.1002/asjc.1362.

[2] J. Delgado-Aguiñaga, O. Begovich, Water leak diagnosis in pressurized pipelines: A real case study, in: C. Verde, L. Torres (Eds.), Modeling and Monitoring of Pipelines and Networks, Vol. 7, Springer International Publishing, 2017, pp. 235–262, http://dx.doi.org/10.1007/978-3-319-55944-5_12.

[3] L. Torres, J. Jiménez-Cabas, O. González, L. Molina, F. López-Estrada, Kalman filters for leak diagnosis in pipelines: Brief history and future research, Journal of Marine Science and Engineering 8 (3) (2020) 173, http://dx.doi.org/10.3390/jmse8030173.

[4] A. Badillo-Olvera, A. Pérez-González, O. Begovich, J. Ruíz-León, Burst detection and localization in water pipelines based on an extended differential evolution algorithm, J. Hydroinform. 21 (4) (2019) 593–606, http://dx.doi.org/10.2166/hydro.2019.123.

[5] A. Pérez-González, A. Badillo-Olvera, O. Begovich, J. Ruíz-León, Calibration zone for the parameters of the differential evolution algorithm and its application to a real burst location problem, J. Fluids Eng. 141 (5) (2019) http://dx.doi.org/10.1115/1.4042749, 051402.

[6] I. Santos-Ruiz, F.-R. López-Estrada, V. Puig, G. Valencia-Palomo, Simultaneous optimal estimation of roughness and minor loss coefficients in a pipeline, Math. Comput. Appl. 25 (3) (2020) 56, http://dx.doi.org/10.3390/mca25030056.

[7] I. Santos-Ruiz, F.-R. López-Estrada, V. Puig, L. Torres, G. Valencia-Palomo, S. Gómez-Peñate, Optimal estimation of the roughness coefficient and friction factor of a pipeline, J. Fluids Eng. (2021) http://dx.doi.org/10.1115/1.4049674.

[8] A. Navarro, J.A. Delgado-Aguiñaga, J.D. Sánchez-Torres, O. Begovich, G. Besançon, Evolutionary observer ensemble for leak diagnosis in water pipelines, Processes 7 (12) (2019) 913, http://dx.doi.org/10.3390/pr7120913.

[9] C. Verde, N. Visairo, S. Gentil, Two leaks isolation in a pipeline by transient response, Adv. Water Resour. 30 (8) (2007) 1711–1721, http://dx.doi.org/10.1016/j.advwatres.2007.01.001.

[10] J. Delgado-Aguiñaga, G. Besançon, O. Begovich, J. Carvajal, Multi-leak diagnosis in pipelines based on extended Kalman filter, Control Eng. Pract. 49 (2016) 139–148, http://dx.doi.org/10.1016/j.conengprac.2015.10.008.

[11] J. Rojas, C. Verde, Adaptive estimation of the hydraulic gradient for the location of multiple leaks in pipelines, Control Eng. Pract. 95 (2020) 104226, http://dx.doi.org/10.1016/j.conengprac.2019.104226.

[12] C. Verde, L. Torres, Referenced model-based observers for locating leaks in a branched pipeline, IFAC-PapersOnLine 48 (21) (2015) 1066–1071, http://dx.doi.org/10.1016/j.ifacol.2015.09.668, 9th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS Paris, 2-4 September 2015..

[13] C. Verde, L. Torres, O. González, Decentralized scheme for leaks' location in a branched pipeline, J. Loss Prev. Process Ind. 43 (2016) 18–28, http://dx.doi.org/10.1016/j.jlp.2016.03.023.

[14] X. Wang, G.A. Camino, T.-C. Che, M.S. Ghidaoui, Factorized wave propagation model in tree-type pipe networks and its application to leak localization, Mech. Syst. Signal Process. 147 (2021) 107116, http://dx.doi.org/10.1016/j.ymssp.2020.107116.

[15] B. Pan, C. Capponi, S. Meniconi, B. Brunone, H.-F. Duan, Efficient leak detection in single and branched polymeric pipeline systems by transient wave analysis, Mech. Syst. Signal Process. 162 (2022) 108084, http://dx.doi.org/10.1016/j.ymssp.2021.108084.

[16] S.H. Kim, Multiple leak detection algorithm for pipe network, Mech. Syst. Signal Process. 139 (2020) 106645, http://dx.doi.org/10.1016/j.ymssp.2020.106645.

[17] J. Delgado-Aguiñaga, G. Besançon, EKF-based leak diagnosis schemes for pipeline networks, IFAC-PapersOnLine 51 (24) (2018) 723–729, http://dx.doi.org/10.1016/j.ifacol.2018.09.655, 10th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS 2018..

[18] G. Besançon, Nonlinear Observers and Applications, Vol. 363, Springer, 2007.

[19] M.H. Chaudhry, Applied Hydraulic Transients, Springer-Verlag New York, 2014.

[20] P.K. Swamee, A.K. Jain, Explicit equations for pipe-flow problems, J. Hydraul. Div. 102 (5) (1976) 657–664.

[21] C. Mataix, Mecánica de Fluidos y Máquinas Hidráulicas, Ed. Del Castillo, 1986.

[22] C.T. Crowe, D.F. Elger, J.A. Roberson, Engineering Fluid Mechanics, John Wiley & Sons, 2000.

[23] E.B. Wylie, V.L. Streeter, Fluid Transients in Systems, Prentice-Hall, 1993.

[24] D. Simon, Optimal State Estimation, John Wiley and Sons, 2006.

[25] J. Gauthier, H. Hammouri, S. Othman, A simple observer for nonlinear systems applications to bioreactors, IEEE Trans. Automat. Control 37 (6) (1992) 875–880, http://dx.doi.org/10.1109/9.256352.

[26] L. Torres, G. Besancon, A. Navarro, O. Begovich, D. Georges, Examples of pipeline monitoring with nonlinear observers and real-data validation, in: 8th International IEEE Multi-Conference on Systems, Signals and Devices (SSD 2011), Sousse, Tunisia, 2011, p. s/n, URL https://hal.archives-ouvertes.fr/hal-00620453.