ORIGINAL ARTICLE

Leak detection and localization in water distribution networks by combining expert knowledge and data-driven models

Adrià Soldevila[1] · Giacomo Boracchi[2] · Manuel Roveri[2] · Sebastian Tornil-Sin[3,4] · Vicenç Puig[3,4]

Abstract
Leaks represent one of the most relevant faults in water distribution networks (WDN), resulting in severe losses. Despite the growing research interest in critical infrastructure monitoring, most of the solutions present in the literature cannot completely address the specific challenges characterizing WDNs, such as the low spatial resolution of measurements (flow and/or pressure recordings) and the scarcity of annotated data. We present a novel integrated solution that addresses these challenges and successfully detects and localizes leaks in WDNs. In particular, we detect leaks by a sequential monitoring algorithm that analyzes the inlet flow, and then we validate each detection by an ad hoc statistical test. We address leak localization as a classification problem, which we can simplify by a customized clustering scheme that gathers locations of the WDN where, due to the low number of sensors, it is not possible to accurately locate leaks. A relevant advantage of the proposed solution is that it exposes interpretable tuning parameters and can integrate knowledge from domain experts to cope with scarcity of annotated data. Experiments, performed on a real dataset of the Barcelona WDN with both real and simulated leaks, show that the proposed solution can improve the leak detection and localization performance with respect to methods proposed in the literature.

Keywords Leak detection · Leak localization · Water distribution networks monitoring · Change detection · Classification

1 Introduction

Water distribution networks (WDNs) are critical infrastructure systems that are difficult to manage and monitor due to their size and complexity. For example, pipes in a WDN of a medium-sized city connect the inlets/reservoirs to hundreds of nodes (either junctions or locations where customers are connected) and span over hundreds of kilometers. In such a large and complex system, faults can be ubiquitous, affecting pipes, reservoirs, sensors or actuators. Leaks, a specific type of hydraulic fault, might occur anywhere as a consequence of pipe breaks, loose joints and fittings, or overflows from storage tanks.

The increasing water demand, pushed by the population growth, and the severe implications of leaks in terms of operational costs and water losses [35], made leak detection and localization a primary concern for water utilities. This has influenced both WDN management strategies and research activities. On the one hand, the vast majority of water management companies nowadays divide the whole

✉ Adrià Soldevila
asoldevila@factic-sf.com

Giacomo Boracchi
giacomo.boracchi@polimi.it

Manuel Roveri
manuel.roveri@polimi.it

Sebastian Tornil-Sin
sebastian.tornil@upc.edu

Vicenç Puig
vicenc.puig@upc.edu

[1] FACTIC, Inc., Urbanització Soldevila, 20, 08650 Cabrianes, Spain

[2] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy

[3] Research Center for Supervision, Safety and Automatic Control (CS2AC), Terrassa, Spain

[4] Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain

WDN into district metered areas (DMAs), where the *flow* and the *pressure* at the inlet can be measured and easily monitored to detect leaks [24, 35]. On the other hand, algorithms for leak detection and localization have been also thoughtfully investigated in control theory [46], computer science [38] and, more recently, artificial intelligence [22, 44]. In particular, most recent solutions monitor recordings from accelerometric sensors [22] or smart meters [26, 54], which provide many measurements and enable sophisticated AI models to be employed. Unfortunately, the vast majority of WDNs are still equipped with flow/pressure sensors at DMA inlets [9, 15, 24, 30, 35, 60], and few flow/pressure sensors deployed inside the DMA.

Despite the promising results achieved by artificial intelligence and neural networks in many domains, leak monitoring remain a challenging problem in particular when analyzing a few flow/pressure recordings (see Sect. 2 for a detailed analysis of the literature) and a general and effective solution is still missing. We speculate the reason is twofold. First, the primary effect of a leak is an anomalous increase in the flow (or a decrease in the pressure), but this is commonly experienced due to variations in the customers' demand, which is difficult to forecast and rarely measured in real-time [16]. Second, although DMAs are typically very large and serve thousands of costumers, these are often equipped with few sensors, because of cost or energy constraints. On top of these critical issues, noise, long-term trends/seasonality, as well as the scarcity of measurements acquired under leak conditions make leak detection and localization very specific and challenging problems requiring ad hoc algorithms. Solutions from related scenarios, e.g., monitoring of a chemical plant or smart grid, do not typically apply [6, 14].

We present a leak detection algorithm that requires only flow measurements at DMA inlets [9, 15, 24, 30, 35, 60], and perform leak localization from a flow/pressure sensors deployed inside the DMA. Our integrated solution comprises three modules: (i) leak detection, (ii) leak validation and leak time/size estimation, (iii) node clustering and leak localization. To compensate for scarcity of sensor information, our algorithms integrate knowledge from domain experts.

We formulate *leak detection and validation* as change-point detection problems, which we solve by an ad hoc two-layer algorithm including a hypothesis test to validate each detection and estimate the leak size and leak time. These latter have been typically ignored by most leak detection algorithms [17, 23, 25, 37, 58, 59], but are crucial to diagnose and localize leaks. Most remarkably, we configure the detection algorithm from few days of flow measurements (without leaks) and from the minimum leak size, a parameter that is easy to interpret and tune for domain experts. We formulate *leak localization* as a

classification problem, and present a solution that is effective even when only a few sensors[1] (e.g., 1 sensor placed per 200 nodes/pipes) acquiring pressure/flow measurements inside the DMA are available. We address leak localization by a set of classifiers that have been specifically trained on sequences generated by a hydraulic simulator of the WDN. Leak localization can seamlessly be trained and used at node level or cluster level [10, 32, 36], where clusters gather nodes where leaks cannot be distinguished, thus allow WDN engineers to set the desired granularity in leak localization. To summarize, we convey the following original contributions:

- A novel leak validation algorithm to reduce false alarms by determining whether each detection corresponds to a sufficiently large leak or not.
- A novel leak localization algorithm, which is based on classifier and is activated every time a detection is validated.
- A specific clustering procedure that gathers nodes where classifiers cannot distinguish the leak location, mainly due to the lack of nearby sensors.

Experiments performed on large datasets of time series acquired in multiple Barcelona DMAs, or that have been simulated from realistic hydraulic models of different cities, demonstrate that the proposed leak-detection and localization algorithms outperform comparable solutions in the literature.

The structure of the paper is the following. Section 2 reviews the literature on leak detection and localization including integrated solutions. Section 3 formulates the leak detection and localization problems, while Sect. 4 gives an overview of the proposed solution. Sections 5 and 6 present in detail the proposed leak detection/validation and leak localization solutions, respectively. Section 7 describes the experiments and discusses results before conclusions that are given in Sect. 8.

2 Related works

In the following, we overview recent leak detection and localization solution with a particular emphasis on those that, like the proposed approach, address both problems with a few flow/pressure sensors are available.

---

[1] Sensor placement is a very important aspect which can heavily influence the localization performance [12], but it is not covered in this work where we assume it has already been done.

2.1 Leak detection techniques

Most leak-detection techniques in the literature monitor the flow measurements at the DMA inlets, which are the most meaningful and always available time series. The mainstream approach consists in (i) fitting a model that well describes the flow time series acquired in leak-free conditions, (ii) computing some residuals or scores between the fitted model and the acquired measurements, and (iii) adopting a statistical/heuristic decision rule to detect leaks.

Several leak-detection algorithms are grounded in statistical or control literature, where models describing the leak-free time series include an adaptive or nonlinear Kalman filter [20, 60]), projections in Fourier domain [15] and particle filters [7]. Data-driven models from AI literature have been used for leak-detection purposes, including support vector regression (SVR) [28], projections over the first principal component analysis (PCA) [30], Bayesian networks [39] and extreme learning machine (ELM) in [43]. Self-similarity of flow time series is instead monitored in [9] thanks to a special feature extraction procedure. In some cases, these models are conveniently used to describe the minimum night flow (MNF), namely the flow during night hours, between 2 am and 6 am, where the flow is minimum and fluctuations w.r.t. patterns are also smoother [35], as illustrated in Fig. 1 with the red dashed line. There are two main reasons for analyzing MNF of the input $F(\cdot)$. First, leaks during MNF are easier to detect as they introduce the largest percentage variation with respect to the total water consumption. Second, the trend of MNF is easier to model, thus any departure from this can be detected as leaks. However, monitoring MNF introduces relevant delays since hours between MNF intervals are not analyzed.

In terms of statistics, most of the above techniques adopt the residuals (possibly normalized or averaged over a time window) between the measurements and model predictions, which are assumed to infer flow in absence of leaks. There is instead more variability in the decision rules adopted, which spans from straightforward thresholding [60], CUmulative SUMmation (CUSUM) test [29] in [20] and the ICI-based change-detection test (CDT) in [9].

None of these algorithms implement specific strategies to mitigate the impact of false alarms that in WDN monitoring are ubiquitous, due to drifts, peaks and seasonality characterizing water consumption. Discarding false alarms is very important, since a high false alarm rate implies relevant economical losses due to unnecessary inspections, and at the same time increases the mistrust of operators to the monitoring system. To this purpose, we customize the hierarchical change-detection framework in [3] introducing a specific validation procedure for flow time series and that

exposes interpretable parameters. In our experiments, we have compared against [9, 15, 30, 60] (described in details in Sect. 7.3) and show that our solution achieves lower detection delays and false negative rates when configured to yield the same false positive rate. Another key advantage of the proposed solution is that we can estimate both the leak starting time and the leak magnitude, which are very important for the localization algorithm but that are rarely provided by competing methods. Our experiments demonstrate that our solution is successful also on real data from the Barcelona WDN, while only a few solutions have been tested on real data [9, 20, 28, 39, 60].

2.2 Leak localization techniques

Many leak-localization algorithms adopt data driven or AI models, and in particular these often resort to training classifiers [36, 47, 49]. Leak localization is typically performed by assuming that a few sensors (most often pressure sensors) have been installed inside the DMA, and that pressure decreases close to the leak. Most solutions in the literature solve this problem by first identifying a candidate region containing many nodes close to the leak, and then pinpointing the exact leak location by inspecting the network using devices such as ground penetrating radars (GPRs) [27]. Various empirical studies [23, 33, 53], localize leaks through mathematical models describing the relation between flow and pressure measurements in presence of leaks. Leak localization can be performed in transient state, using a model of [11] the dynamic effects of the leak in the time series like negative pressure waves, or in steady-state, namely comparing the flow and pressure measurements inside the DMA against a reference that was acquired/generated/modeled in absence of leaks. Steady-state methods are the most popular ones, since they typically require fewer sensors than transient state ones. A few steady-state methods employ correlation analysis [32], k-NN [47] or more powerful classifiers [36, 49] that take as input residuals between measurements inside the DMA and the output of an hydraulic simulator of the WDN. In our experiments, we compared against [32, 47, 49], which we better describe in Sect. 7.4. Among the aforementioned works, only [27, 32, 47, 49] were validated on real data.

Our leak localization algorithm also relies on classifiers that take as input a richer descriptor of the WDN status than [32, 47, 49]. In particular, to cope with leaks of different magnitudes, we train a collection of classifiers: one per each expected leak size, and every time we select which one to use. To train this model at best, we resort to data-augmentation procedures inspired by [13], and here expanded. Moreover, the proposed leak localization is coupled with a clustering algorithm to group locations where leaks are more difficult to localize. The granularity
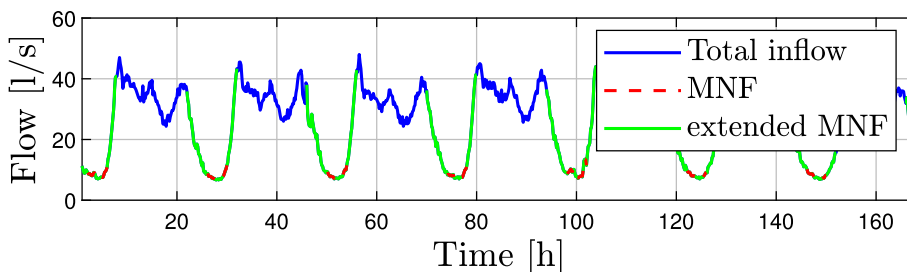
Fig. 1 Example of weekly profile for the total DMA inflow $F(t)$. The MNF period spans from 2 a.m. to 6 a.m. and is highlighted in red while the extended MNF period spans from 10 p.m. to 8 a.m. and is highlighted in green. This picture is better interpreted in the colored version of the paper (color figure online)

of clustering results can be easily adjusted by experts, thus representing a very useful tool to monitor DMA equipped with few sensors. Most remarkably, once clustering is adopted, classifiers are seamlessly retrained and used at cluster level. Previous solutions [47] adopt clustering as a post-processing phase and not as a joint step to be combined with classifiers used for localization.

2.3 Integrated solutions

A few integrated solutions that perform both leak detection and localization have been presented [17, 59], which, however, require a large number of special sensors operating at high sampling rates. As such, these solutions are not easy to adopt in most DMA. Fuzzy theory has also been used for simultaneously detecting and localizing leaks [19, 55]. These solutions address the different forms of uncertainties characterizing WDNs, such as nodal demand variability and sensor noise, but without any validation step. Another relevant similarity to ours solution is that the parameters regulating different fuzzy states are interpretable (e.g., leaks size) and can be defined by domain experts. This approach has also been recently pursued in [57] to monitor and control smart homes.

3 Problem statement

3.1 Leak detection

We consider the leak-detection problem by monitoring the *total inflow* $F(\cdot)$ of a DMA[2], which is a time series sampled at regular time intervals, that in absence of leaks measures the amount of water $\Psi(t)$ consumed within the DMA each time instant t. For this reason, $F(\cdot)$ exhibits a repetitive pattern on a daily basis [5], which depends on weekends, holidays or weather conditions, as shown in Fig. 1. A leak permanently modifies the flow $F(t)$ by increasing the water

---

[2] On DMAs provided with multiple inlets, $F(\cdot)$ sums the flow measured in all these

consumption of an unknown leak size $l > 0$ at the leak-starting time T^*, namely:

$$F(t) = \begin{cases} \Psi(t), & t < T^* \\ \Psi(t) + l, & t \geq T^* \end{cases}. \qquad (1)$$

We assume that the leak size is constant. Even though leaks often gradually increase over time, this approximation typically holds over short time intervals, where the leak has to be detected [10, 32].

Our primary goal is to detect when a leak occurs inside a DMA and accordingly estimate both the leak-starting time T^* and size l. The detection time, i.e., the time instant when a monitoring algorithm reports a leak, is denoted as $\widehat{T}$, while $\widehat{T}^*$ and $\widehat{l}$ denotes the estimated leak-time and the estimated leak size, respectively. We assume that only leaks above a minimum size $l_{\min}$ need to be reported. A good detection algorithm should provide short *detection delays* (DD) $\widehat{T} - T^*$, and very low *false negatives rate* (FNR), namely the percentage of leaks above the minimum size $l_{\min}$ that have not been detected. At the same time, the *false positive rate* (FPR), namely is the percentage of detections where there is not a leak should be kept as low as possible. It is further assumed that leak detection algorithm has to be configured from a training sequence H containing the first days of flow measurements without leaks (Fig. 2).

3.2 Leak localization

We also consider the leak localization problem, which consists in estimating, after each detection at $\widehat{T}$, the node j^* where the leak has occurred. To this purpose, we assume that a few pressure/flow sensors have been deployed inside the DMA and that the i-th node records either the time series of pressure $p_i(\cdot)$ or the flow $f_{i,j}(\cdot)$ between nodes i and j.

We assume that sensors inside the network are very sparse, i.e., that in total there are only m time series recorded and that $m << n$, being n the total number of
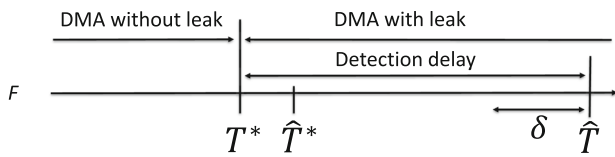
Fig. 2 Leak detection notation: leak time T^*, estimated leak time $\widehat{T}^*$, leak detection time $\widehat{T}$, δ corresponds to the number of values used for leak validation

candidate leak locations. Another typical assumption in the literature we make is that there are no simultaneous leaks in different locations [32]. The estimated leak location $\widehat{\jmath}$ has to be as close as possible to the true leak location j^*, where the distance can be either measured in terms of pipe length, node or linear distance.

Gathering a representative training set for leak localization purposes is unfeasible in the real world, as this would require measurements of flow and pressure in all the n possible leak locations and for different leak sizes. Hence, we simulate a training set **TR** of flow/pressure time series at nodes inside the DMA. To this purpose, we require: (i) a training sequence of leak-free inlet flows like the one used for leak detection, (ii) the time series of leak-free recordings from the m internal measurements, (iii) a calibrated hydraulic model of the DMA and (iv) a base nodal demands ξ, i.e., the percentage of water consumed by each node (even based on monthly bills).

4 An overview of the proposed solution

Figure 3 illustrates the proposed solution, which comprises three main modules: (i) the *leak detection* module, (ii) the *leak validation* module and (iii) the *leak localization* module. The *leak detection* module monitors the total inflow $F(\cdot)$ at DMA inlets by means of a change-detection test that compares the acquired data w.r.t. leak-free flow measurements. Once a change has been detected at time $\widehat{T}$, the change-detection test also provides an estimate of the leak starting time $\widehat{T}^*$, which is used to activate the *leak validation* module. Validation module further analyzes the flow at inlets to reduce false positive detections by means of an ad hoc statistical hypothesis test comparing the leak-free flow measurements with the measurements acquired between $\widehat{T}$ and $\widehat{T}^*$.

When the detection is confirmed, the leak size $\widehat{l}$ is estimated by comparing the flow time series before and after the estimated leak time $\widehat{T}^*$. Domain experts play a crucial role in the validation module, as they can set the minimum leak size $l_{\min}$ to be detected, and this greatly contributes to discarding false alarms and detections due to fluctuations and other non-stationarities in the flow time series. We emphasize that both the detection and validation algorithms require a short training set of leak-free measurements from the total inflow time series $F(\cdot)$. The leak detection and validation modules are described in detail in Sect. 5.

Once a leak has been detected, validated and the leak size $\widehat{l}$ estimated, the *leak localization* module is triggered, which analyzes measurements from the m sensors placed inside the DMA to estimate the leak location—denoted by $\widehat{\jmath}$. To achieve this goal, the leak localization module relies on a set of classifiers trained on synthetically generated time series, which encompasses leaks in each of the n considered locations and for different leak sizes. All these time series are generated by means of the hydraulic model of the network, which is fed to a *simulator* (as, e.g., Epanet [41]) together with historical leak-free flow recordings and data-augmentation guidelines provided by domain *experts*. During training, an iterative spectral-clustering algorithm operating with the expert-in-the-loop, aggregates nodes where classifiers would not be able to localize leaks, to carry out localization at the level of clusters rather than nodes. The leak localization module is described in detail in Sect. 6.

5 Leak detection and validation

Instead of pursuing the common approach of monitoring the MNF of the inflow F (see Sect. 2), we monitor the total inflow during the *extended minimum night flow* (eMNF), which covers a longer period where still the flow exhibits controlled variations. Figure 1 compares the MNF and the eMNF over a week and shows that eMNF includes the MNF. We define the eMNF E time series as a portion of F (i.e., $E \subset F$) that spans everyday between 10 p.m. and 8 a.m. for the residential areas we consider in our experiments. When the DMA serves industrial areas, this period must be accordingly set by domain experts. Even though we exclude from eMNF high demand hours (as these would require a very long training set to distinguish fluctuations due to customer's demand or leaks), monitoring eMNF requires a more general and flexible model than MNF.

Leaks can be conveniently detected by monitoring eMNF time series through change-detection tests (CDTs) [8], which are sequential techniques to detect even negligible—but persistent—changes in a data generating process. Unfortunately, the vast majority of CDTs in the statistical literature apply only to data streams composed of independent and identically distributed (i.i.d.) realizations of a random variable. This is not the case of the flow F, nor E, that instead are time series showing repeated patterns on a daily basis (see Fig. 1). This type of regularity can be
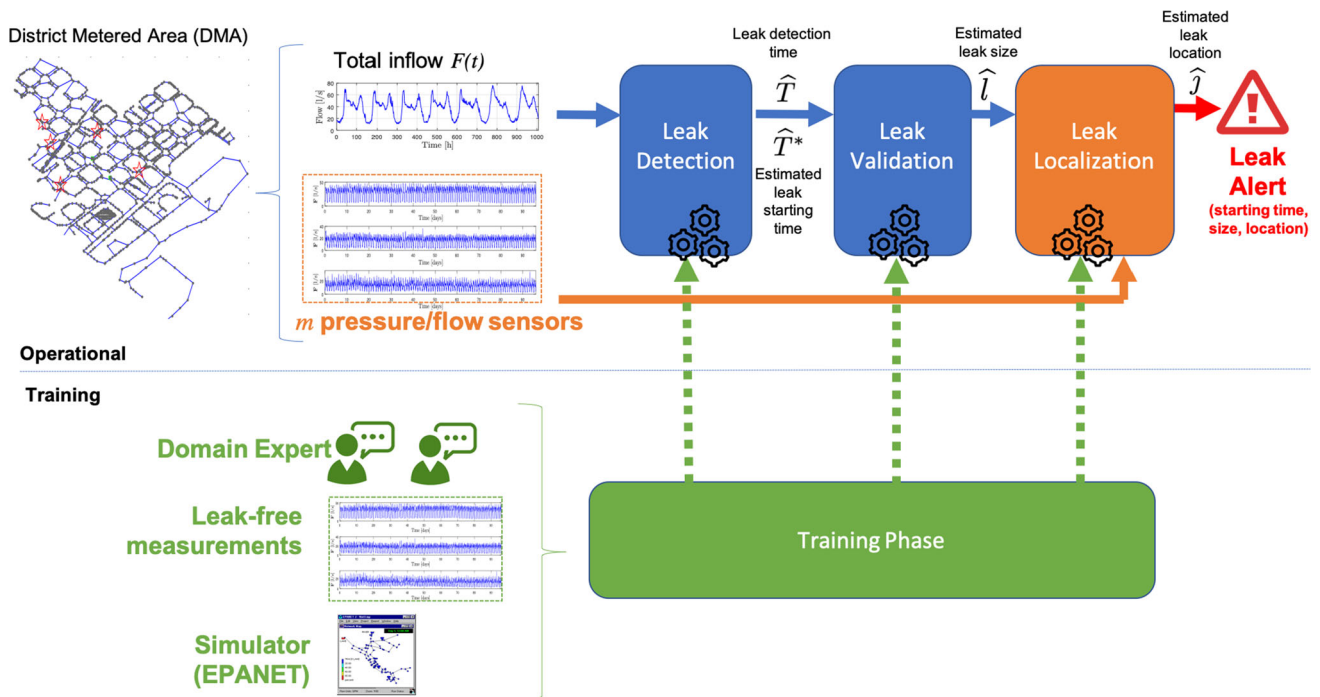
Fig. 3 An overview of the proposed solution comprising leak detection, leak validation and leak localization. All these modules are trained from a flow time series and take advantage of knowledge from domain experts. This picture is better interpreted in the colored version of the paper (color figure online)

enforced as in [9] to extract a sequence of features values ϱ that assess the similarity of an input time series with a reference leak-free training sequence. Thus, we can successfully monitor E by a CDT analyzing a stream of i.i.d. realizations from an unknown random variable.

We expect the distribution of features ϱ to change when a leak occurs. However, distribution changes might also occur as a consequence of abnormal demands, seasonal drifts or sensor errors, to name a few examples. To prevent these common situations from raising an unacceptable number of false alarms, we implement the hierarchical change-detection test formulation proposed in [4], and we designed a monitoring scheme composed of two modules (illustrated in Fig. 4) specifically meant for leak-detection purposes. Our first module performs the feature extraction and monitoring of ϱ by a sequential CDT. While there are no strict limitations on the CDT to be employed, this has to reveal even subtle changes in the daily consumption patterns: such variations, when persistent, might indicate a leak. Our second module determines whether the prospective leak affects the monitored DMA in a realistic manner, and to this purpose we analyze the flow measurements directly. In what follows, we provide a detailed description of the proposed hierarchical CDT for leak detection.

5.1 Feature extraction and change detection

We extract ϱ features to assess whether each small patch of incoming flow measurements is similar to those in the training set as in [9]. A *patch* $\mathbf{s}_t$ is a short sequence extracted from the eMNF, namely:

$$\mathbf{s}_t = \{E(t - v), \ldots, E(t), \ldots, E(t + v)\}, \tag{2}$$

where the time t represents the patch center, and v is the number of samples selected on each side of the patch, such that the patch size is $2v + 1$. We compute features ϱ by comparing patches extracted from the input flow time series against patches extracted from the first q days of the initial training sequence, namely H_q. Thus, $H_q \subset H \subset E$ and this is recorded under leak-free conditions. For each input patch $\mathbf{s}_t$, the closest patch in H_q in terms of Euclidean distance to $\mathbf{s}_t$ is selected among those referring to the same time of the day. We denote $\boldsymbol{\pi}_t$ as the most similar patch to $\mathbf{s}_t$ among the training ones belonging to H_q:

$$\boldsymbol{\pi}_t = \underset{\xi}{\arg\min} \|\mathbf{s}_t - \mathbf{s}_\xi\|_2, \tag{3}$$

where the minimization is performed over patches having centers $\xi \in \{h(t), \beta + h(t), 2\beta + h(t), \ldots\}$, being $h(t)$ the time of the day associated to t, and $\beta = 24$ h denotes the daily cycle characterizing the monitored time series. Thus, the most similar patch $\boldsymbol{\pi}_t$ is selected from H_q, as long as
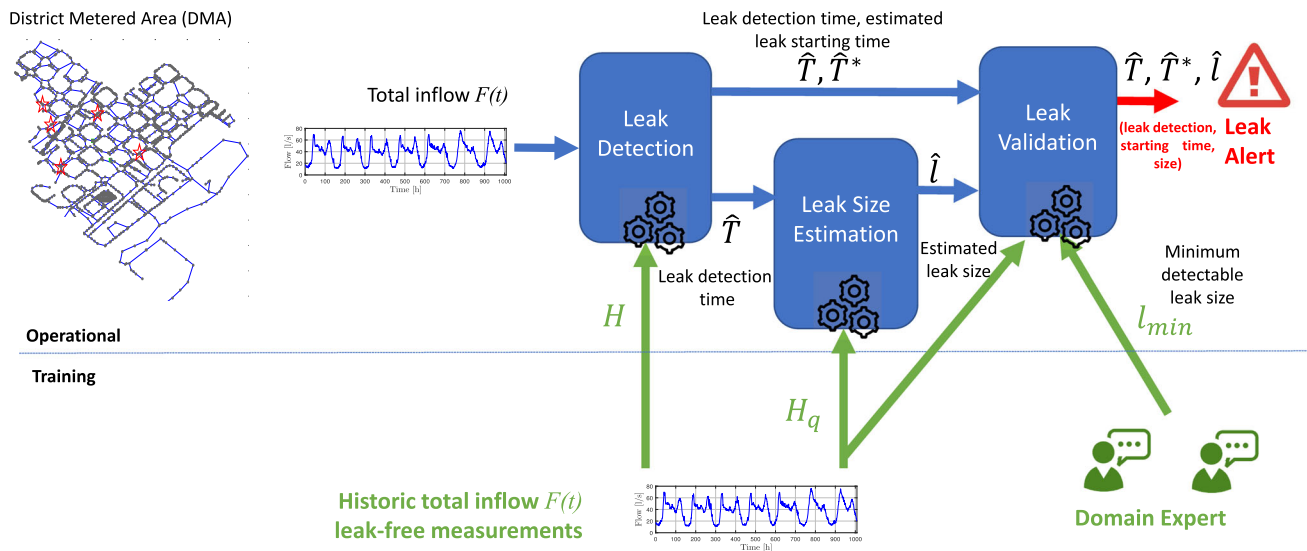
Fig. 4 Leak detection scheme, where historical leak-free flow measurements at the inlets are used as reference for leak detection and leak size estimation, and the information of those modules is used in combination with domain experts knowledge setting the minimum leak size l_{min} to validate the leak. This picture is better interpreted in the colored version of the paper (color figure online)

this refers to the same time of the day as $\mathbf{s}_t$. In (3), $\| \cdot \|_2$ denotes the ℓ_2 norm of a vector. The feature $\varrho(t)$ is defined as the difference between the center of $\mathbf{s}_t$ and the center of $\boldsymbol{\pi}_t$ denoted by (2):

$$\varrho(t) = \mathbf{s}_t(v+1) - \boldsymbol{\pi}_t(v+1). \tag{4}$$

As discussed in [9] and tested for the specific case of flow time series, the ϱ values can be approximated as i.i.d. realizations of a random variable, thus can be monitored by most CDTs. Similarly to [9], we adopt the intersection-of-confidence-interval (ICI)-based CDT [1], which monitors ϱ over disjoint windows. In particular, this test first computes the sample mean and a power-low transformation of the sample variance (to approach a Gaussian distribution) over each incoming window. These values are then used to update the global estimates of the same quantities over the entire sequence. These global estimates (which are assumed to be constant in the change-detection framework) are analyzed together with their confidence interval to detect distribution changes. The amplitude of these confidence intervals is defined by the tuning parameter Γ, which regulates the CDT promptness in detecting changes. More precisely, the ICI rule [18] detects a change in ϱ as soon as the intersection of all the intervals from these global estimates becomes empty. The CDT requires only a portion of ϱ time series for configuration, and these have to be extracted from training patches that are not in H_q. Therefore, we configure the ICI-based CDT form features extracted from H_r, namely the remaining r days in $H = [H_q, H_r]$, being $[\cdot, \cdot]$ the time series concatenation. Further details on the ICI-based CDT can be found in [1].

The CDT at the first module detects any change affecting either the mean or the variance of ϱ, which can be in principle due to a non-leak event. This is the reason why we designed the following validation module for detected leaks.

5.2 Validation

To reduce the FPR, each detected change has to be confirmed by the validation module (Fig. 4), which assesses whether there is evidence of a leak in the specific DMA. To this end, we adopt *i*) a paired one-sided Wilcoxon's test [56], which is a hypothesis test meant to determine whether the median of an unknown distribution has changed, and *ii*) we define at each DMA, l_{min} the size of the smallest leak that is expected to be detected. Typically, WDN engineers employed in the monitoring can define a suitable value of l_{min}, which often ranges between 5% and 10% of the average inflow.

We define E_{TS} as a vector representing the average daily inflow over H_q, i.e., during the first q training days:

$$E_{TS}(h(t)) = \frac{1}{q} \sum_{i=0}^{q-1} H_q(h(t) + i\beta), \tag{5}$$

where $h(t)$ is the position of t in the current day and β is defined as in (3). An example of E_{TS} is depicted in Fig. 5. After each detection, the latest $\delta > 0$ measurements preceding $\widehat{T}$ are selected, i.e., $\{E(\widehat{T} - \delta), \ldots, E(\widehat{T})\}$ and we remove any trend in the eMNF, by computing the point-wise difference between a window of the same size opened over recent data and E_{TS}:
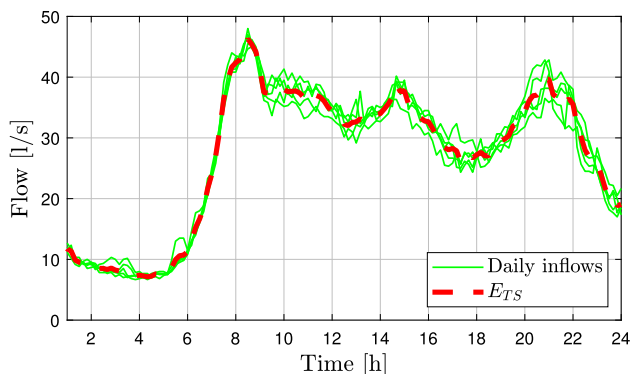
Fig. 5 Example of the E_{TS} vector obtained by averaging from five daily samples from training set H_q

$$\triangle E(i) = E(\widehat{T} - \delta + i) - E_{TS}\left(h(\widehat{T} - \delta + i)\right) - l_{\min,} \qquad (6)$$
$$\text{for } i = 1, \ldots, \delta.$$

Note that in the right hand side of (6) we subtract $l_{\min}$ to validate only leaks larger than the minimum leak size. We then validate leaks by running a paired and one-sided Wilcoxon's test [56] with confidence level α over $\triangle E$, thus determining whether there is enough statistical evidence for claiming that (6) is above zero, thus there is a leak larger than $l_{\min}$.

Every time the null hypothesis is rejected, the detection is validated, and thus, we activate the leak localization module. To this purpose, we first estimate the leak size:

$$\widehat{l} = \frac{1}{\delta} \sum_{t=\widehat{T}-\delta}^{\widehat{T}-1} \left(E(t) - E_{TS}(h(t))\right). \qquad (7)$$

The change time can be estimated by the ICI-based CDT [3] through a retrospective analysis after each detection. A few other change-detection algorithms, like the change point method (CPM) [40], provide such an estimate after each detection. On top of leak localization, which is the primary task for WDN utilities, it is also possible to activate heuristic procedures for re-training/adapting the CDT as commented in [2]. In WDN monitoring, these heuristics might be useful for compensating variations in the customers demand.

When there is not enough statistical evidence to reject the null hypothesis, we discard the detection and all the data before $\widehat{T}$. In particular, the CDT at the leak detection module is returned to monitor the inflow at time $\widehat{T} + 1$.

It is worth mentioning that, to compute $\triangle E$ in (6), it might be necessary to manipulate sequences to compensate for seasonal drifts. This is in particular feasible when two DMAs exhibiting similar behavior are simultaneously being monitored, and the trend estimated from one sequence can be used to detrend the other.

6 Leak localization

Our leak localization module is illustrated in Fig. 6, and comprises a set of classifiers which have been specifically designed for localizing the leak inside the DMA. Each classifier processes flow and pressure measurements acquired inside the DMA and predicts the leak location (Sect. 6.1). Since the leak size influences much the input time series, we train a set of classifiers $\{C'_l\}$ each one corresponding to a leak size l, which is a parameter varying in a predefined range.

The most critical aspect of our supervised learning approach is the shortage of training data. In fact we would need, for each considered leak size l, measurements affected by a leak in each and every network location, which of course is not a viable option. Therefore, like other works in leak localization literature [32, 49], we adopt an hydraulic model of the DMA and a *simulator* (e.g., Epanet [41]), together with historical leak-free flow recordings estimated water demands from customers to generate a large set of flow/pressure time series referring to nodes inside the DMA. Here, domain *experts* play a primary role in defining data-augmentation guidelines and transformations that manipulate the flow time series and customer demand (Sect. 6.2) to yield a realistic training set **TR**.

During the training phase, we can aggregate nodes where classifiers would not be able to exactly localize leak, by the clustering algorithm proposed in Sect. 6.3. This is an iterative spectral clustering procedure, which takes into account classifiers previously trained to assess how accurately a leak can be detected. Domain experts play a central role during clustering as well, since they might visualize clusters being created during iterations, and stop the process at the desired level of granularity. Leaks are then conveniently localized at cluster level (some of which might also consist of a single node) and the same localization algorithm can be seamlessly used, after training.

6.1 Leak identification by classification

In what follows we define the classifiers $\{C_l\}$ and, for the sake of notation, we omit the leak size l where this is not necessary. We train each classifier C to analyze the flow and pressure measurements inside the DMA and determine where the leak has occurred among the n possible locations. Measurements at DMA inlets are not informative enough to locate leaks; therefore, we require that $m << n$ sensors (either flow or pressure ones) were placed inside the DMA at known locations.

After each validated detection, we quantitatively assess the impact of the leak inside the DMA by averaging the
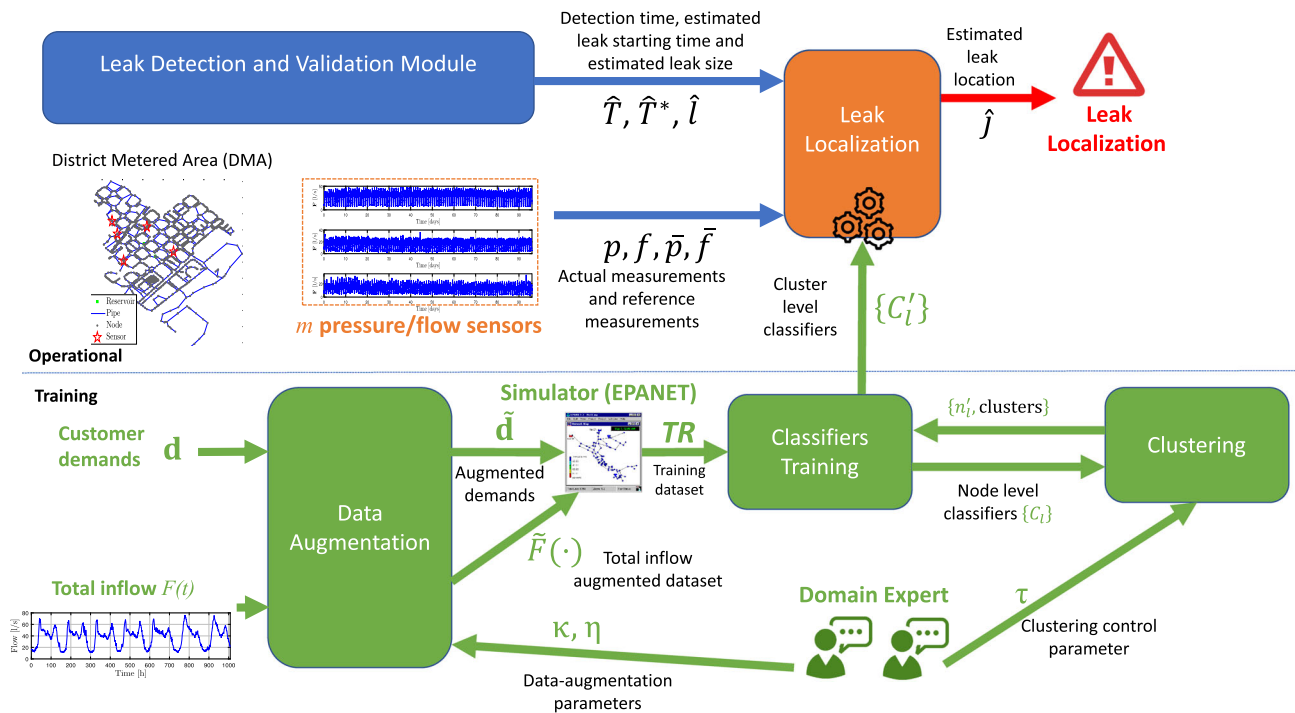
Fig. 6 Leak localization scheme. Historical measurements from inlet sensors and nodal demands from billing records are augmented by procedures defined by domain *experts*, and then used to generate a new training set **TR** with the hydraulic simulator of the network. **TR** is used to train the node-level classifiers, which are then used in the clustering phase to group the locations where leak localization is not possible. Cluster-level classifiers $\{\mathcal{C}'_l\}$ are then trained and used for inference (during the operational phase). This picture is better interpreted in the colored version of the paper (color figure online)

variations at these sensors before and after the estimated leak-time $\widehat{T}^*$:

$$\triangle p_i = \frac{1}{\left(\widehat{T} - \widehat{T}^* + 1\right)} \sum_{t=\widehat{T}^*}^{\widehat{T}} \left(p_i(t) - \overline{p}_i(t)\right)$$

$$\triangle f_{i,j} = \frac{1}{\left(\widehat{T} - \widehat{T}^* + 1\right)} \sum_{t=\widehat{T}^*}^{\widehat{T}} \left(f_{i,j}(t) - \overline{f}_{i,j}(t)\right),$$

(8)

where p_i denotes the pressure measurements acquired at the ith node, while $f_{i,j}$ denotes the flow measurements between nodes i and j. The terms $\overline{p}_i$ and $\overline{f}_{i,j}$ denote reference measurements recorded without leaks in the same location. Similarly to E_{TS} in (5) and Fig. 5, we compute $\overline{p}_i$ (resp. $\overline{f}_{i,j}$) by averaging measurements over different days in the training time series acquired at the m internal sensors during H_q as in (5). For leak localization purposes, differences are computed by aligning p_i (resp. $f_{i,j}$) and $\overline{p}_i$ (resp. $\overline{f}_{i,j}$) at the same time of the day. Note that in (8) we do not consider the eMNF period, but rather the entire time series.

We define the input **x** of a classifier $\mathcal{C}$, as a m-dimensional vector having in each component the variation in either flow or pressure due to the leak as in (8):

$$\mathbf{x} = [\triangle\mathbf{f}, \quad \triangle\mathbf{p}]^{\mathrm{T}}, \, x \in \mathbb{R}^m. \tag{9}$$

We train the classifier $\mathcal{C}$ to provide as output the correct leak location among the n nodes of the DMA. Thus, the estimated leak location is:

$$\widehat{\jmath} = \mathcal{C}(\mathbf{x}), \; \widehat{\jmath} \in \{1, \ldots, n\}. \tag{10}$$

In particular, we train a maximum likelihood classifier $\mathcal{C}$ that builds upon class-specific density models. Thus, we associate to each potential leak location $j \in \{1, \ldots, n\}$ a m-dimensional Gaussian density model $\Phi_j = \mathcal{N}(\mu_j, \Sigma_j)$, where $\mu_j \in \mathbb{R}^m, \Sigma_j \in \mathbb{R}^{m \times m}$. The choice of the Gaussian distribution is rather customary in the leak localization literature [1, 48, 49] and, at the same time, ease the node clustering procedure described in Sect. 6.3. Thus, for each input sample **x**, we compute $\Phi_j(\mathbf{x})$ for each class $j \in \{1, \ldots, n\}$ and associate **x** to the class $\widehat{\jmath}$ yielding the largest posterior probability by means of:

$$\widehat{\jmath} = \mathcal{C}(\mathbf{x}) = \underset{j \in \{1, \ldots, n\}}{\operatorname{argmax}} \left(\log\left(\Phi_j(\mathbf{x})\right)\right). \tag{11}$$

The parameters of the classifier $\mathcal{C}$ are n pairs (μ_j, Σ_j) $j = 1, \ldots, n$, which describe each density Φ_j. These parameters are obtained by sample estimators computed

from a synthetic training set **TR** obtained through simulation as discussed in what follows.

We emphasize that $\mathcal{C}$ depends on the leak size l, as this can completely change the input **x**. Therefore, z different values of leak magnitude l are considered, resulting in z different classifiers $\{\mathcal{C}_l\} = \{\mathcal{C}_1, \ldots, \mathcal{C}_z\}$ trained. During operations, the classifier associated with the leak size that best matches $\widehat{l}$ estimated during leak detection is selected. As discussed in the following, it is unfeasible to acquire a training set for each of these classifiers. Thus, these training sets are generated through a specific data-augmentation procedure that uses the hydraulic model of the DMA.

6.2 Data-augmentation and training set preparation

As mentioned before, we generate multiple leak-free sequences of flow and pressure measurements by means of the Epanet simulator [41]. This is fed with realistic time series of inlet flow $\tilde{F}$ and customer demands $\tilde{d}_i$, $\{i = 1, \ldots, n\}$, which are obtained by a data-augmentation procedure that was agreed with domain *experts*. The procedure is depicted in the bottom part of Fig. 6.

Each augmented total inflow $\tilde{F}$ is obtained from F as:

$$\tilde{F}(t) = F(t + \lambda) + \kappa(t), \tag{12}$$

where λ is a small random time-shift, and κ is a term that can be either zero or defined to modify a portion of $F(t + \lambda)$. In particular, κ can introduce a few spikes or replace a portion of $F(t + \lambda)$ with another measurement recorded in the same hour in a different day.

The augmented demand at the i-th node $\tilde{d}_i$ is defined from historical billing records as in [34]. In particular, we first infer from the historical billing records all the *base-demands* $\{\xi_i\}$, where $\xi_i \in [0, 1]$ is the portion of the total inlet flow F that reaches the i-th node. As a consequence, base demands sums to one $\sum_{i=1}^n \xi_i = 1$. We simulate a time series from each nodal demand by adding a time-variant uncertainty $\eta(t)$ term over the expected value ξ_i, as in [13]:

$$\tilde{\xi}_i(t) = \xi_i + \eta(t), \quad i = 1, \ldots, n, \tag{13}$$

where $\eta(\cdot)$ is white Gaussian noise $\mathcal{N}(0, 0.25)$ truncated in $[-0.5, 0.5]$. We then obtain the data-augmented nodal demands as follows:

$$\tilde{d}_i(t) = \frac{\tilde{\xi}_i(t)}{\sum_{i=1}^n \tilde{\xi}_i(t)} \tilde{F}(t). \tag{14}$$

The augmented nodal demand at the i-th node $\tilde{d}_i(t)$ is thus proportional to augmented total inflow $\tilde{F}(t)$ and to the percentage of augmented nodal demand, which has been

rescaled to sum to 1 in each time instant t. Division by $\sum_{i=1}^n \tilde{\xi}_i(t)$ performs such rescaling.

We generate leaks of size l at node i by introducing a steady extra demand at the specific location i:

$$\tilde{d}_i^{(l)}(t) = \frac{\tilde{\xi}_i(t)}{\sum_{i=1}^n \tilde{\xi}_i(t)} (\tilde{F}(t) - l) + l. \tag{15}$$

In contrast, in any location without leak $j \neq i$, we adjust the nodal demands as:

$$\tilde{d}_j^{(l)}(t) = \frac{\tilde{\xi}_i(t)}{\sum_{i=1}^n \tilde{\xi}_i(t)} (\tilde{F}(t) - l), j \neq i. \tag{16}$$

This is a rather common practice in WDN monitoring [10, 32], and corresponds to first subtracting the leak amount l from the total inflow $\tilde{F}$, and then adding the leak amount l exclusively to the time series of the selected leak location i.

Time series of augmented demands before and after the leak ($\{\tilde{d}_i\}$ and $\{\tilde{d}_i^{(l)}\}$, respectively) are fed to the Epanet simulation to generate flow $\{f_{i,j}\}$ and pressure time series $\{p_i\}$ inside the DMA. The same procedure is repeated for multiple values of the leak size l and leak locations $i = \{1, \ldots, n\}$.

We further manipulate flow $\{f_{i,j}\}$ and pressure time series $\{p_i\}$—either with or without leaks—by introducing a multiplicative random term to mimic sensor noise:

$$\tilde{f}_{i,j}(t) = f_{i,j}(t)(1 + \eta(t)), \tag{17}$$

where $\eta(\cdot)$ is white Gaussian noise $\mathcal{N}(0, 0.25)$ truncated in $[-0.5, 0.5]$ as in (13) to add larger uncertainty where the flow is larger. Augmented pressure measurements $\tilde{p}_i$ are generated in a similar way.

Both augmented flows $\tilde{f}_{i,j}$ and pressure $\tilde{p}_i$ time series are then used to train the classifier as in Sect. 6.1. In particular, Fig. 7 summarizes the adopted procedure to artificially generate training sequences. Each complete sequence consists of an initial part without leak, followed by a second part containing a leak of l [l/s], introduced as an extra demand as in (15) and (16). These two are then used as in (8) to generate the features needed to train the classifiers. This procedure is repeated for each potential leak location $\widehat{j} = 1, \ldots, n$, and for each leak size considered to yield a meaningful training set for the classifiers in (11).

6.3 Clustering nodes for leak localization

The large uncertainty on nodal demands makes leak localization a very challenging problem, thus leak localization estimates can be very poor when the number of sensors inside the DMA is small. In particular, in the regions of the classifier's input space where Gaussians Φ_j
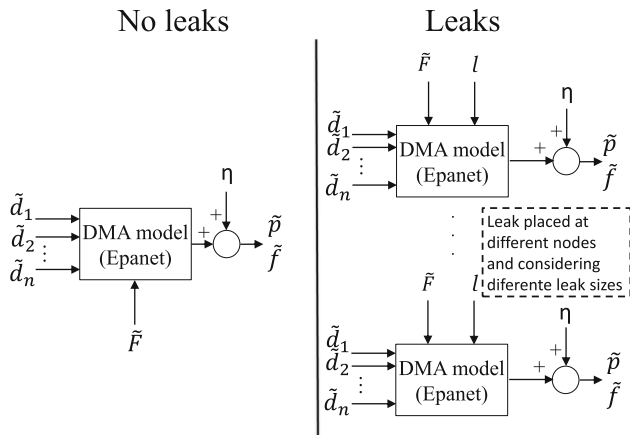
No leaks　　　Leaks

Fig. 7 Scheme of the procedure used for data generation using a hydraulic simulator. Different sequences are generated for training our classifiers where the first part is leak free while the second part introduces a leak at different nodes and for each considered leak size

largely overlap, it might not be possible to exactly locate leaks. Thus, we propose an algorithm to cluster nodes and map the localization uncertainty over the DMA layout. This clustering can help WDN engineers to identify those regions where leaks cannot be exactly pinpointed, and localization should be performed at cluster level rather than at node level.

We formulate node clustering in a DMA as a cut problem on a weighted undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ similar to [22, 44]. Each graph vertex $\mathcal{V}$ corresponds to one of the n candidate leak locations and each edge $\mathcal{E}$ corresponds to a pipe connecting two nodes. Clustering is solved by an iterative algorithm, the graph cuts [45]. The graph initially associated with a DMA contains a single connected component, since all the nodes are reached by the total flow from inlets. The graph-cut algorithm performs a recursive splitting of the graph, where the sub-graphs are the results of cuts that minimize an energy functional. Splits are determined by the eigenvalues of the weight matrix $\mathbf{W}$ of the graph, and the process is terminated by standard stopping criteria, like the functional value, the maximum number of calls and the minimum number of vertices in sub-graphs.

The weight matrix $\mathbf{W}$ is a $n \times n$ matrix where each row and column corresponds to a candidate leak location. To effectively solve leak localization, the weight matrix $\mathbf{W}$ has to be defined—for each DMA and classifier $\mathcal{C}$ – upon a specific distance measure. The weight associated to two directly connected nodes i and j is defined as:

$$\mathbf{W}_{i,j} = e^{-\left(\mathrm{sKL}(\Phi_i, \Phi_j)/\tau\right)^2}, \tag{18}$$

where $\mathrm{sKL}\left(\Phi_i, \Phi_j\right)$ denotes the symmetric Kullback–Leibler (sKL) divergence and τ is a user defined parameter to control the clustering. When $\mathrm{sKL}(\Phi_i, \Phi_j) = 1$ nodes i and j

are very distinguishable, while $\mathrm{sKL}(\Phi_i, \Phi_j) = 0$ corresponds to nodes that are not distinguishable. The $\mathrm{sKL}(\Phi_i, \Phi_j)$ is defined as $\mathrm{sKL}(\Phi_i, \Phi_j) = \frac{1}{2}(\mathrm{KL}(\Phi_i, \Phi_j) + \mathrm{KL}(\Phi_j, \Phi_i))$, and is a distance measure between distributions that range in $[0, 1]$. In case of Gaussian functions, $\mathrm{KL}(\Phi_i, \Phi_j)$ can be computed through a closed form expression:

$$\mathrm{KL}(\Phi_i, \Phi_j) = \frac{1}{2}\left(\mathrm{tr}(\Sigma_j^{-1}\Sigma_i) + (\mu_j - \mu_i)^T\Sigma_j^{-1}(\mu_j - \mu_i) - m \right.$$
$$\left. + \ln\left(\frac{\det(\Sigma_j)}{\det(\Sigma_i)}\right)\right), \tag{19}$$

where $\mathrm{tr}(\cdot)$ denotes the trace and $\det(\cdot)$ the determinant of a matrix and m is the dimension of the space where distributions Φ_i, Φ_j lives. The parameter τ in (18) controls how fast the node distance increases with the sKL. This is a special parameter of graph cuts, which has to be set by domain experts that might take into account the number of sensors and the magnitude of the input flow (we experienced smaller τ are preferable when flow is large) or following the procedure in Section 3.1 of [45].

As shown in Fig. 6, once this iterative splitting procedure is terminated, each sub-graph represents a cluster of nodes where leaks are not distinguishable, except from sub-graphs containing a single node. The number of clusters corresponds to the number of locations we denote by n', where leaks can be located. Once nodes are aggregated in clusters, node-level classifiers $\mathcal{C}$ have to be replaced by cluster-level classifiers by computing the Gaussian densities $\{\Phi'\}$ over each non-singleton cluster. This corresponds to running the same procedure described in Sect. 6.1, and yields a new classifier $\mathcal{C}'$ operating at cluster level, thus returning values in $1, \ldots, n'$.

Note that since the weight matrix in (18) is defined depending on a specific classifier $\mathcal{C}$ trained at node level, the whole clustering procedure needs to be run for each of the leak sizes considered in the set of classifiers $\{\mathcal{C}_l\}$. The set $\{\mathcal{C}'_l\}$ corresponds to all the retrained classifiers operating at cluster level for different leak sizes. Once trained, the set $\{\mathcal{C}'_l\}$ is fed to the leak localization module, which selects the classifier corresponding to the estimated leak size.

Since the stopping criteria for graph cuts are rather arbitrary and dictated by practical arguments, it is useful to display the sub-graphs created at each iteration, and let WDN engineers choose the best level of clustering. This also allows the identification of the most challenging regions of the DMA for leak-localization purposes.

7 Experiments

We test our solution in three real-world case studies, where this is compared against solutions widely used in the leak detection and leak localization literature. More precisely, we assess leak detection performance over real measurements from five DMAs from the Barcelona WDN where leaks have been artificially introduced. We test the integrated leak detection and localization solution in artificial data from the Limassol DMA, and in a real leak scenario from the Nova Icària DMA in Barcelona.

7.1 Figures of merit

We adopt several figures of merit from the pattern recognition literature [42, 50–52] to assess the leak detection and localization performance.

7.1.1 Leak detection and size estimation

We consider the following indicators to evaluate the performance of the proposed leak detection and leak-size estimation methods, which are computed over all the sequences during eMNF hours:

- FPR or false positive rate is the percentage of sequences having a false detection, thus a leak detected at time $\widehat{T} < T^*$.
- FNR or false negative rate is the percentage of leaks that have not been detected.
- DD or detection delay is the difference between the true leak starting time and the detection time as $\widehat{T} - T^*$, expressed in hours and considering the entire day/night, not just eMNF.
- DTD or difference time detection is the difference between the true leak starting time and the estimated leak starting time as $\widehat{T}^* - T^*$, expressed in hours like DD.
- The average error in the leak size estimation $\triangle \widehat{l}$ expressed in [l/s].

We emphasize that DD, DTD and $\triangle \widehat{l}$ are computed only on correct leak detections.

7.1.2 Leak localization

We assess leak localization performance as the accuracy indicator χ and its modified version ω, which takes into account the fact that localization occurs at cluster level. These indicators are obtained from the confusion matrix $\mathbf{\Upsilon}$ that is commonly used in classification. Every entry $\Upsilon_{i,j}$ corresponds to the number of leaks at node i that have been located in node j. A perfect classification would yield to a

diagonal $\mathbf{\Upsilon}$. The overall adjusted accuracy ω is expressed as:

$$\omega = 100 \frac{\sum_{i=1}^{n'} \Upsilon_{i,i} \frac{1}{u_i}}{\sum_{i=1}^{n'} \sum_{j=1}^{n'} \Upsilon_{i,j}}, \tag{20}$$

where u_i is the number of nodes in the ith cluster and n' the number of clusters. This is meant to measure classification performance at cluster level. When no clustering is performed or when all the clusters result in singletons, this indicator is replaced by χ, i.e., the percentage of correctly localized leaks defined as:

$$\chi = 100 \frac{\sum_{i=1}^{n} \Upsilon_{i,i}}{\sum_{i=1}^{n} \sum_{j=1}^{n} \Upsilon_{i,j}}. \tag{21}$$

Note that an ideal algorithm should achieve both $\chi=100$ and $\omega=100$.

7.2 Configuration of the proposed solution

We configure the ICI-based CDT by setting $\Gamma = 1$ and $v = 6$, such that patches contain 13 samples. The Wilcoxon's test at the validation layer was configured with $\alpha = 0.05$ and has been executed over a window δ opened over the past 6 h (the actual value of δ therefore depends on the sampling rate as these can be 36 or 72 samples in the considered case studies). The value of $l_{\min}$ in the validation layer was selected depending on the DMA characteristics, and the same for the clustering parameter τ: the values of these parameters are summarized in Table 1. We emphasize that the proposed techniques have been compared against widely used leak detection and localization methods described in Sects. 7.3 (leak detection) and 7.4 (leak localization), respectively.

7.3 Leak detection methods for comparison

We compare the proposed solution against the following leak detection algorithms. To enable a fair comparison, all these techniques have been configured over the same training set to yield, or at least approach where not possible, the same FPR value.

7.3.1 ICI-based CDT (ICI-CDT)

This is the same technique used at the detection layer [9], without validation layer. Therefore, this requires setting $\Gamma = 4.6$ to achieve the same FPR in Barcelona DMAs and $\Gamma = 2$ in the other two case studies. Other tuning parameters are set the same as in the proposed solution. This method has been considered to assess the improvement provided by the proposed validation layer.

Table 1 Characteristics of the case studies experiments for small, medium, large and real leak sizes

DMA	$\bar{F}$	#Res.	#PRVs	n	#Pipes	Small	Medium	Large	l_r	q	r	l_{min}	τ
Bellamar	6.7	2	0	1523	154	1	2	4	-	10	4	0.5	–
Gavà Centre (set 1)	18.2	2	0	3373	3482	2	3	5	–	10	4	0.5	–
Can Roca	10.0	1	0	1427	1473	1	2	4	–	10	4	0.5	–
Parc de la M. (set 1)	37.3	2	0	1507	1553	3	5	8	–	10	4	0.5	–
Gavà Centre (set 2)	16.1	2	0	3373	3482	2	3	5	–	10	4	0.5	–
Canyars	24.6	1	0	692	717	2	3	5	–	10	4	0.5	–
Parc de la M. (set 2)	14.8	2	0	1507	1553	2	3	5	–	10	4	0.5	–
Limassol	0.14	1	1	47	58	0.125	0.250	0.375	–	10	2	0.05	10
Nova Icària	50.9	2	2	1520	1646	–	–	–	5.6	2	0.5	3.8	5

#Res. is the number of reservoirs, #PRVs the number of PRVs, n the number of potential leak locations, #Pipes the number of pipes, q the number of patterns for the self-similarity, r the number of patterns for the ICI CDT, l_{min} the minimum leak size expected to be detectable and τ the user-defined threshold for clustering

7.3.2 Leak detection based on PCA (LD-PCA)

This method, proposed in [30], relies on dimensionality reduction to jointly analyze multiple flow measurements. Here, all the flow measurements over one day are stacked in a vector (where each attribute is a flow measurement) and then vectors for multiple days are stacked in a matrix. This is done for both recent measurements to be analyzed and historical ones that are leak free. Then, the PCA transformation of the historical matrix is computed and the loads covering at least 95% of the variance are selected. The same number of principal components is selected from the matrix of recent measurements and the extracted loads are compared. A leak is detected when the difference in loads exceeds a certain threshold. Other approaches use statistical features extracted from current and past measurements. To guarantee the same FPR as other methods, we set the threshold as the mean value of the loads plus 3.7 times the standard deviation computed over the training set for Barcelona DMAs (1.1 times the standard deviation in the Limassol DMA). Due to the limited amount of data provided for training, it has not been possible to configure this method for the Nova Icària leak case.

7.3.3 Adaptive Kalman filter (AKF)

This method, introduced in [60], relies on a Kalman filter to predict the flow and generate normalized residuals for each recording in a week. Normalized residuals are then averaged over a sliding window spanning 1 week, and compared against a threshold to detect a leak. Here, the threshold was set to 0.19 in the Barcelona DMAs, while it has not been possible to tune the method to achieve the same FPR in the Limassol DMA. Our intuition is that this is due to large fluctuations on the water consumption pattern probably caused by the small number of customers. Therefore, we adopt the same threshold as for the Barcelona DMAs. Finally, the threshold is set to 0.05 in the Nova Icària DMA.

7.3.4 CUSUM test for Fourier coefficients (Fourier-CUSUM)

This solution [15] relies on the first Fourier coefficient on a window opened over the past, leak-free, measurements to normalize the inlet flow. The same normalization is applied to the incoming measurements and the first Fourier coefficient is compared against a threshold. The same work presents an alternative approach using the same normalization, but leaks are detected when the maximum difference with the most similar flow pattern in the last few days persistently exceeds a threshold. The latter approach has been adopted in this experimental section. To achieve the target FPR, we set a threshold to 0.38 for the Barcelona DMAs and 0.59 Limassol DMA and we required two consecutive days of detections (namely days where the residuals exceed the threshold, instead of to one in [15]). Finally, in Nova Icària DMA the threshold is set to 0.13, while the minimum number of consecutive days of detections is set to zero.

7.4 Considered leak localization methods

We compare against three techniques following a steady-state approach:

7.4.1 Leak-signature correlation (LS-Corr)

This solution, presented in [32], relies on a hydraulic simulator to estimate the pressure, and then computes residuals w.r.t. the recorded measurements. Residuals are

then compared against the sensitivity matrix (which is computed off-line and contains the expected residuals for each leak location and size) and the node having the highest correlation is selected as the leak node candidate. We configure this method like our localization solution, but over residuals computed using simulations, and without adding noise or other demand uncertainties during data augmentation. Residuals are computed hourly, yielding many leak-location estimates that are combined over time according to [32].

7.4.2 k-nearest neighbor (k-NN)

This solution, introduced in [47], relies on the same residual computation as in [32] but it also integrates demand and noise uncertainties during data-augmentation for training the model. We selected $k=3$ in the k-NN. As in the previous method, the residuals are computed hourly and aggregated over time as the authors suggested.

7.4.3 Bayesian reasoning (Bayesian)

This approach, suggested in [49], relies on the computation of residuals and the assumption that each potential leak location fits a Gaussian distribution in the feature space. The classifier is used as suggested in [49] considering residuals with uncertainty and a time horizon aggregation.

In all the above methods, classifiers were configured using residuals as in [47, 49] on sequences generated with the same data-augmentation procedure. It should be highlighted that the two latter classifiers do not require leak size as an additional input. Other classifiers have not been considered since these typically achieve comparable performance, as shown in [36].

7.5 Leak detection in Barcelona DMAs

We exhaustively tested the proposed leak detection procedure on five different DMAs of Barcelona WDN. The main characteristics of these DMAs (numbers of reservoirs, pressure reducing valves (PRVs), nodes and pipes) are summarized in Table 1. In all the DMAs, the flow at inlets has been recorded in leak-free conditions over two time periods: from January 1st 2013 until May 18th 2013, and from August 31st 2013 to March 3rd 2014, which correspond to set 1 and set 2 in Table 1, respectively. Days affected by missing values or outliers (values three times larger than the flow mean) have been removed from the two sets, resulting in 85 days from the first set, and accordingly 85 days from the second. For each DMA, we assemble three sequences of flow at inlets spanning days 1-55, 16-70, 31-85, obtaining overall 21 sequences among all DMAs and the two sets. In each sequence, the first 14

days are used for training, the next 21 are without leaks and in the remaining 20 days, a leak has been synthetically injected as described in (1). The data sampling rate in all these DMAs is 10 min.

Three different leak magnitudes are considered: *small*, *medium* and *large* leaks whose magnitude depends on the average total inflow and summarized in Table 1. Since the eMNF (10 p.m.–8 a.m.) is considered, every time the leak is not detected before the 8 a.m., the detection is delayed at least 14 h.

Detection results are summarized in Table 2. The proposed leak detection technique outperforms the alternative on small leaks, and in particular, it is the most successful in terms of FNR. This is a very important aspect considering that false negatives would correspond to a substantial increase in the DD when large testing sequences were provided. In terms of medium and large leaks, the AKF and the Fourier-CUSUM solutions achieve slightly better performance, having lower DD and FNR. Compared to ICI-CDT, the proposed solution is prompter at detecting changes, thanks to the validation module it can be configured with a lower value of Γ yielding the same FPR. The proposed solution provides instead more accurate estimates of leak time and size. Note that, due to the eMNF time interval considered, it is rather easy to achieve large detection delays.

7.6 Leak detection and localization in limassol DMA

We consider the DMA of Limassol WDN (Fig. 8) as a second case study to test the whole integrated leak detection and localization solution. Out of the 57 consumer nodes, only those that are located downstream a pressure reducing valve (PRV) are considered as potential leak locations, which results in $n=47$ nodes. In our simulations, we assume that two pressure sensors have been installed in nodes 16 and 28, and that the pipes between nodes 16 and 18 and between 27 and 28 are equipped with flow sensors. Other details of this case study are summarized in Table 1.

7.6.1 Leak detection

The inlet flow time series used for leak detection has a sampling rate of 5 min and lasts 130 days. The first 10 days of measurements were used to generate the training set of classifiers by data-augmentation as illustrated in Sect. 6. Out of the remaining 120 days, 5 sequences (corresponding to day intervals 1–48, 19–66, 37–84, 55–102 and 73–120 days) composed by 12 days for training, 18 days without leak and 18 days with leak are generated considering three different leak sizes.

Table 2 Leak detection performance in Barcelona and Limassol DMAs

Method	Small leaks					Medium leaks					Large leaks				
	FPR	FNR	DD	DTD	$\Delta\hat{l}$	FPR	FNR	DD	DTD	$\Delta\hat{l}$	FPR	FNR	DD	DTD	$\Delta\hat{l}$
Barcelona DMAs case study															
Proposed	9.5	9.5	123.9	76.4	0.8	9.5	4.8	73.9	28.6	0.8	9.5	0	49.2	3.9	1.2
			(75.3)	(69.3)				(31.5)	(31.2)				(18.1)	(17.7)	
ICI-CDT	9.5	57.1	221.7	195.6	–	9.5	23.8	268.6	221.4	–	9.5	0	122.7	72.1	–
			(175.9)	(176.2)				(173.6)	(180.2)				(95.0)	(94.1)	
LD-PCA	9.5	33.3	124.7	–	–	9.5	23.8	73.7	–	–	9.5	4.8	48	–	–
			(164.4)					(45.6)					(41.0)		
AKF	9.5	47.6	146.0	–	4.2	9.5	4.8	29.4	–	2.6	9.5	0	25.5	–	2.2
			(80.2)					(88.9)					(68.9)		
Fourier-CUSUM	9.5	28.6	74.5	–	–	9.5	0	72.4	–	–	9.5	0	72.0	–	–
			(39.7)					(24.6)					(24.0)		
Limassol DMA case study															
Proposed	20	0	24.6	1.5	0.03	20	0	5.8	0.5	0.05	20	0	5.9	0.6	0.07
			(12.1)	(16.1)				(11.2)	(15.2)				(11.2)	(15.2)	
ICI-CDT	20	0	24.9	2.2	–	20	0	6.0	0.6	–	20	0	22.8	0.9	–
			(11.8)	(14.6)				(9.3)	(13.3)				(9.5)	(13.5)	
LD-PCA	20	0	32.0	–	–	20	0	32.0	–	–	20	0	32.0	–	–
			(1152.0)					(1152.0)					(1159.0)		
AKF	100	100	–	–	–	100	100	-	–	–	100	100	-	–	–
			(–)					(-)					(-)		
Fourier-CUSUM	20	0	52.4	–	–	20	0	52.4	–	–	20	0	52.4	–	–
			(26.6)					(26.6)					(26.6)		

Mean and standard deviation represented in parenthesis are shown. To enable a fair comparison, these methods were configured to achieve the same FPR, when possible. Tests on Barcelona DMAs contain 21 sequences, while Limassol DMA only 5, hence the coarse granularity of false positive rate
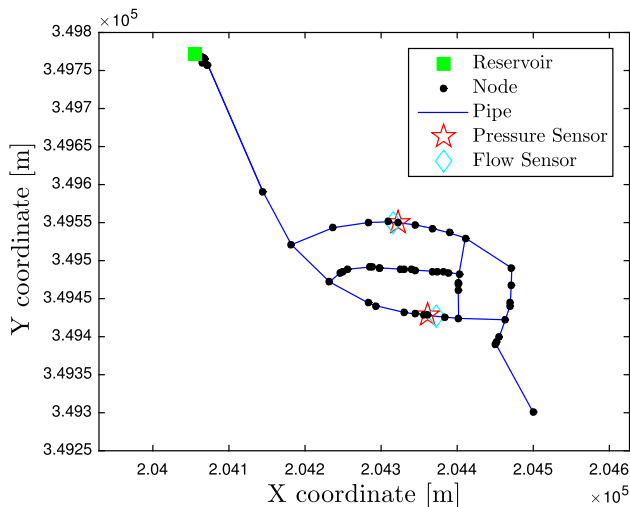

Fig. 8 Topology of Limassol DMA

As in the previous case study, leaks are artificially injected and three different sizes are considered: small leaks, with size 0.125 [l/s], medium leaks, with size 0.250

[l/s] and large leaks, with size 0.375 [l/s]. The mean value for the input flow is 0.14 [l/s]. Since the total inflow and the sampling frequency are different from the Barcelona DMAs, the detection layer has been tuned as in the Barcelona DMAs, while the validation layer was tuned as follows: l_{min}=0.05 [l/s] and a δ=72 measurements (corresponding to 6 h recordings).

The leak detection performance is reported in Table 2 and confirms that the proposed solution outperforms all the others in terms of FNR. Note that, since there are only five sequences, a single false positive results in a 20% FPR. This is why we were not able to achieve 10% FPR as in the Barcelona DMAs case. Nevertheless, all the methods were configured to achieve 20% FPR, except from AKF as discussed in Sect. 7.3.

7.6.2 Leak localization

We adopt the data-augmentation procedure described in Sect. 6.2 to generate time series for testing leak localization. The augmented procedures been configured as

follows: the first 10 days of inlet measurements are used to generate 50 sequences as in (12). Both real and augmented sequences have been fed to the Epanet hydraulic simulator to generate the flow and pressure measurements at sensors placed inside the network, where nodal demands have been modified as in Eqs. (13, 14, 15, 16). These measurements are used to generate the training set to estimate the parameters of the Gaussian distributions and perform clustering. We configure the clustering process by setting $\tau = 10$ in (18), and in each iteration we split the graph in a number of subgraphs corresponding to the number of smallest eigenvalues having their cumulative sum below 0.05. However, in each iteration, we enable a maximum number of 5 splits.

The results of the clustering procedure described for classifiers C^l trained in medium leak size are depicted in Figs. 9 and 10 for pressure and flow sensors, respectively. Using pressure sensors, 14 non-singleton clusters with maximum of 5 nodes are obtained. This is because the low consumption results in small variations of the pressure inside the network, preventing to distinguish the location of the leak with pressure sensors, resulting in larger clusters. Using flow sensors, only 3 non-singletons were formed with a maximum of three nodes. Flow sensors are better able to distinguish the leak, since the leak flow represents a relevant portion of the total inflow. Since flow is more heavily affected than pressure, leaks are expected to be easier to locate, thus fewer clusters appear in clustering driven by flow.

Leak localization performance is summarized in Table 3 and shows that the proposed leak-localization algorithm performs particularly well in the Limassol DMA in the case
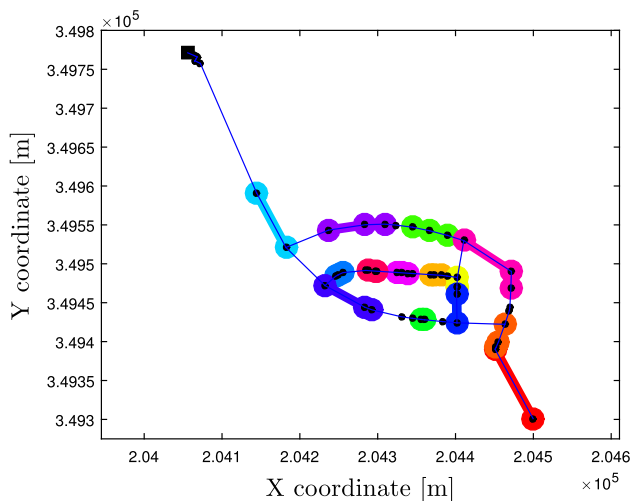

Fig. 10 Clusters formed using classifiers C^l trained on flow sensors and medium leak. Each color represents a different cluster. Singleton nodes of the clusters are black. This picture is better interpreted in the colored version of the paper (color figure online)
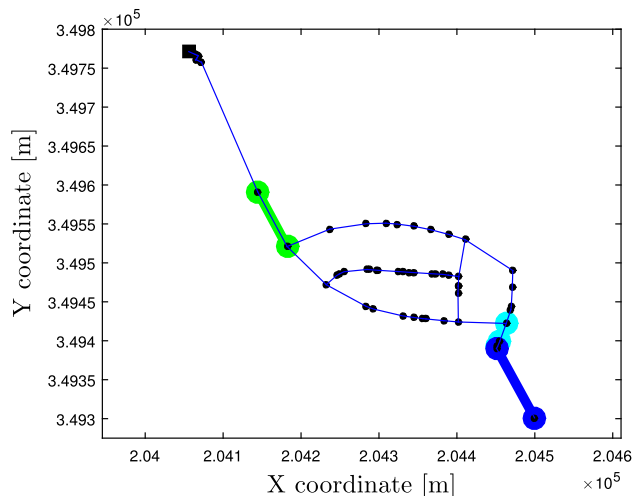
of flow sensors, in particular for medium leaks and large leaks where the results combined with clustering delivers very precise localization. This is not the case of other techniques, that are not able to distinguish leaks at different locations. Perhaps, the problem lies in the nature of the residuals that these techniques use: different leak locations are typically very overlapped in the residual space [47], yielding very poor localization performance. Localization performance using pressure sensors is very poor and this is because pressure falls due to the leaks are not very

Table 3 Leak localization results in Limassol DMA, considering that three different leaks at 47 nodes

Technique	Ind.	Sens.	Small leaks	Medium leaks	Large leaks
Without clustering					
Proposed	χ	p.	2.1	2.7	3.2
	χ	f.	37.8	50.0	54.8
LS-Corr	χ	p.	2.7	2.1	1.1
	χ	f.	6.9	4.8	2.1
k-NN	χ	p.	2.1	2.1	2.1
	χ	f.	5.9	4.3	4.3
Bayesian	χ	p.	0	0	0
	χ	f.	2.1	2.1	2.1
With clustering					
Proposed	χ	p.	6.6	3.6	2.5
	ω	p.	5.3	2.8	1.0
	χ	f.	43.5	82.7	84.2
	ω	f.	40.3	80.1	81.8

p and *f* in the "Sens." column stand for pressure and flow sensors. The larger these indicators, the better

Fig. 9 Clusters formed using classifiers C^l trained on pressure sensors and medium leak. Each color represents a different cluster. Singleton nodes of the clusters are black. This picture is better interpreted in the colored version of the paper (color figure online)

noticeable, as will be discussed in Sect. 7.8. In this case, also the proposed clustering solution does not improve much the localization performance.

7.7 Leak detection and localization in Nova Icària DMA real case

The third and final case study is entirely based on real measurements acquired in the Nova Icària DMA, another DMA of Barcelona WDN. This DMA has two reservoirs with flow measurements and PRVs. Inside the DMA, five pressure sensors are placed in nodes 3, 4, 5, 6 and 7 using the methodology described in [31]. The topology of the network and the sensor placement are depicted in Fig. 11. Most relevant network parameters are in Table 1, which indicates the small number of sensors employed (5) compared to the large number of candidate leak locations (1520).

Differently from the two previous cases, these measurements have been recorded in a real leak scenario: acquired data contain six days of flow measurements without leaks, 30 h of data with leak and another 16 h without leak. The leak was introduced by opening a fire hydrant by the company in charge of the network management, resulting in a leak size of approximately 5.6 [l/s]. The configuration parameters used in the detection and localization procedures are the same as in the previous case studies. The only difference is the minimum leak size, which was set to 5% of the average consumption of water, namely $l_{\min}$=3.8 [l/s].

Figure 12 shows the results of the proposed leak detection technique: the first plot presents inlet flow, the second the eMNF used in the leak detection procedure, and the last one the extracted features ϱ: the blue line indicates the training set for extracting features, the orange line the

training set for the ICI-based CDT, the red line indicates the leak time T^*, magenta line indicates the estimated leak time $\widehat{T}^*$, and green line indicates the detection time $\widehat{T}$. The detection has a delay of 177 samples from the whole sequence (not only eMNF) corresponding to 29.5 h. The difference between T^* and $\widehat{T}^*$ is only one sample, i.e., 10 min. The estimated leak size is 7.1 [l/s] with respect to 5.6 [l/s] in reality. The ICI-CDT delivers the same results, while the LD-PCA method was not able to detect the leak, probably due to the short training set available. The application of AKF was very successful, detecting the leak with a delay of five samples (40 min), estimating a leak size of 4.3 [l/s]. Finally, the Fourier-CUSUM technique detected the leak after 10 h (one complete period of eMNF).

The proposed clustering algorithm applied on the recordings from the five pressure sensors (configured using τ=5) provides clusters depicted in Fig. 13. The small number of sensors employed makes hard to distinguish the location of the leak at each node, but still it can be appreciated a superior performance achieved by pressure sensors than in Limassol DMA. We speculate that this is probably due to the larger water consumption. It can be noticed that the resulting clusters are very consistent with the information spread through the network, since singletons are close to the sensor's nodes, while far from these, the clusters definitively increase their size.

Table 4 reports the leak localization performance computed in two different settings. The first one "After detected the leak" consists in activating the proposed leak localization algorithm in cascade to the leak detection. Localization is configured from the estimated $\widehat{T}^*$ and $\widehat{l}$. In this case, the proposed algorithm localizes the leak in the singleton cluster at the node 474, while the real leak is at
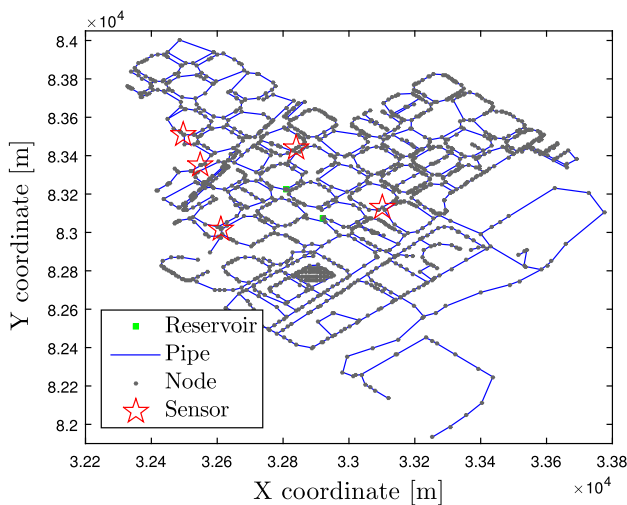

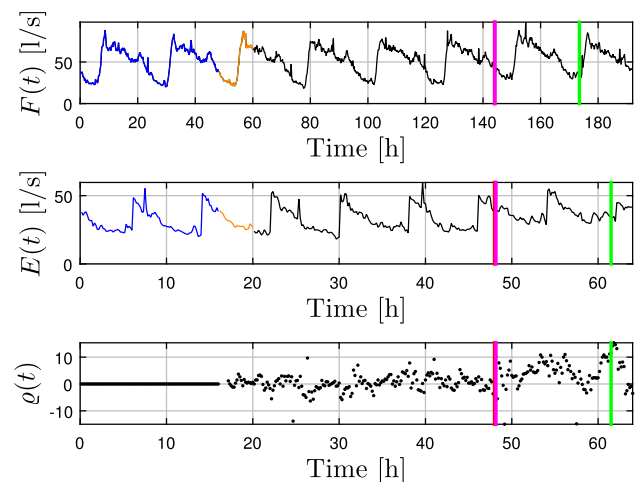**Fig. 11** Topology of Nova Icària DMA

Fig. 12 Leak detection for Nova Icària DMA. This picture is better interpreted in the colored version of the paper (color figure online)
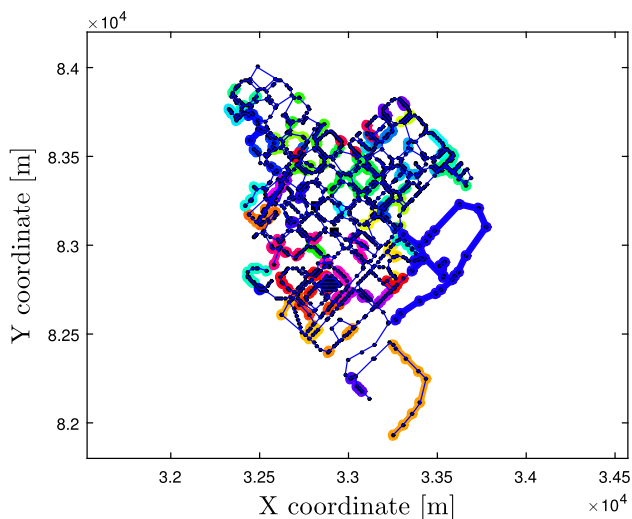
Fig. 13 Nova Icària clustering results. Clustering is performed over the five pressure sensors and considering a leak of 5.6 [l/s]. Colors indicate different non-singleton clusters. This picture is better interpreted in the colored version of the paper (color figure online)
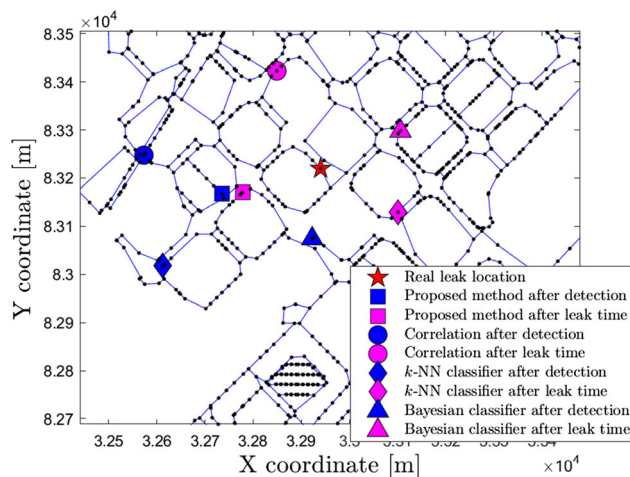

Fig. 14 Nova Icària leak localization results. Blue markers corresponds to localizations made in the "After detected the leak" scenario, while magenta markers correspond localizations in the "After true leak time" scenario. This picture is better interpreted in the colored version of the paper (color figure online)

Table 4 Leak localization results in the Nova Icària real case

Technique	After detected the leak			After true leak time		
	Linear dis.	Pipe dis.	Node dis.	Linear dis.	Pipe dis.	Node dis.
Proposed	211.9	452.9	21	169.4	422.1	20
LS-Corr	367.8	478.6	26	222.0	433.4	17
k-NN	384.7	667.9	12	184.0	390.8	13
Bayesian	146.6	932.4	33	183.2	265.0	10

Linear distance and pipe distance are measured in meters, while the nodes distance is measured in nodes. The smaller the better for these indicators

node 996. Other leak localization algorithms [32, 47, 49] result in different node candidates: 1508 for the LS-Corr method, 4 for the *k*-NN and 1 for the Bayesian. Their relative locations are shown in Fig. 14. The second scenario is referred to as "*After true time*", and it assumes that the leak is perfectly detected such that 24 h of leaky data are provided. These are the same settings as in [32, 47, 49]. In this case, the proposed method returns a singleton cluster, the node 1463 (see Fig. 14) along with the localization performance presented in [32, 47, 49] for other methods. Table 4 summarizes some indicators used to assess localization performance, and it can be seen that the proposed technique delivers better results in terms of pipe distance—which is the most meaningful one for water companies that pinpoint the leak by searching pipe by pipe—in the most realistic settings where leak localization is performed in cascade to a leak-detection algorithm.

7.8 Discussion

The proposed leak-detection algorithm outperforms other solutions in both Barcelona and Limassol case studies at least for small leaks. This also achieves equivalent or superior performance in terms of FNR for medium and large leaks. The proposed validation layer always improves the leak-detection performance and this is in agreement with previous findings in change detection [3]. Hopefully, this suggests that the validation module can provide a performance boost also in combination with other leak-detection techniques. In fact, thanks to the validation module, the ICI-CDT can afford configurations that are more prompt in detecting changes, considering that false positives are filtered by the validation module.

Regarding the proposed leak localization solution, it has shown to perform well when monitoring features extracted from flow measurements as in (9). Analyzing pressure seems instead not effective in the Limassol DMA, while it enables a very accurate localization performance in the Nova Icària DMA. This is due to the different topology and

hydraulic condition of the two networks. As discussed in [21], the larger the flow in nominal conditions, the higher the relative impact of a leak of a given size would be on the pressure measurements. Thus, considering that the overall flow values in the Limassol DMA is rather small, the pressure fall due to a leak should be almost negligible for the considered leak sizes. This is the reason why leak localization performance—when analyzing pressure measurements—are very poor for all the considered techniques. In the Nova Icària DMA, the flow is larger in leak-free conditions, and the few pressures sensors employed are able to sense the leak. The proposed solution in the realistic scenario, where it is configured from the estimates provided by the leak detection algorithm, can localize leaks with the lowest pipe distance.

8 Conclusions

In this paper, we proposed a comprehensive leak monitoring solution for WDNs, which wisely combines machine learning models and information coming from domain experts to tackle the challenging problems of WDN monitoring. The proposed method covers both leak detection and localization tasks in an integrated manner. In particular, the proposed ad hoc validation module is used in cascade with the detection module and allows detecting subtle leaks leading to a reduced FNR and DD. Leak detection has been proven effective on real data with real and injected leaks, outperforming other methods. The experiments also demonstrate that monitoring the eMNF is particularly effective in small networks, where the daily patterns are subject to large fluctuations compared to their standard consumption. Also, the proposed leak detection algorithm yields very reliable estimates of the leak time and size, which are used by the leak localization algorithm.

The proposed leak localization algorithm is entirely data-driven, and requires only a hydraulic model of the network to generate a meaningful training set by data-augmentation. The leak localization algorithm is very accurate, and outperforms all the competing methods in combination characterized by flow sensors in networks with low water consumption. Our experiments indicate that analyzing flow and pressure differences before and after the estimated leak starting time yields superior localization performance than directly classifying residuals, which is the mainstream approach in the literature [32, 47]. Finally, we present a algorithm to cluster nodes where leaks cannot be distinguished, which turns also in an inspection method to identify regions of the DMA where leak localization is too difficult due to the limited number of sensors installed inside the network.

Our solution shares a few limitations of other methods in the literature as it detects and localizes one leak at a time and it assumes that leaks occur at nodes only. Moreover, the number of the sensors deployed inside the network and their locations are key for an effective leak localization. This is particularly relevant for branches of the WDN without sensors, where leak localization might not be possible. Another relevant aspect influencing both detection and localization performance is the service pressure, which can increase the leak size and make pressure drop more apparent. Despite these limitations, we have shown that our solution successfully combines machine learning methods and knowledge from WDN engineers, e.g., for setting the minimum leak size to be detected and in the clustering procedure.

Finally, although our proposed solution addresses leak detection and localization in WDNs, the general methodology and the key ideas on which it is based (leak detection improved with leak validation, leak localization based on distributed measurements and classifiers, dataset augmentation using simulation, use of clustering procedure to gather nodes where classification is not possible, role of domain experts to tune the solution) have the potential to be applied to other types of critical large-scale distribution networks such as oil and gas networks.

Availability of data and materials Leak detection time series in Barcelona DMAs can be found in https://boracchi.faculty.polimi.it/Projects/SelfSimilarityCDT.html while we do not share Limassol and Nova Icària data that are under non-disclosure agreement.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Alippi C, Boracchi G, Roveri M (2010) Change detection tests using the ICI rule. In: The 2010 international joint conference on neural networks (IJCNN), pp 1 –7
2. Alippi C, Boracchi G, Roveri M (2016) Hierarchical change-detection tests. IEEE Trans Neural Netw Learn Syst PP(99):1–13

3. Alippi C, Boracchi G, Roveri M (2017) Hierarchical change-detection tests. IEEE Trans Neural Netw Learn Syst 28(2):246–258

4. Alippi C, Scotti F (2006) Exploiting application locality to design low-complexity, highly performing, and power-aware embedded classifiers. IEEE Trans Neural Netw 17(3):745–754

5. Alvisi S, Franchini M, Marinelli A (2007) A short-term, pattern-based model for water-demand forecasting. J Hydroinf 9(1):39–50

6. Amiruddin AAAM, Zabiri H, Taqvi SAA, Tufa LD (2020) Neural network applications in fault diagnosis and detection: an overview of implementations in engineering-related systems. Neural Comput Appl 56:1–26

7. Anjana GR, Sheetal-Kumar KR, Mohan-Kumar MS, Amrutur B (2015) A particle filter based leak detection technique for water distribution systems. Procedia Eng 119:28–34

8. Basseville M, Nikiforov IV (1993) Detection of abrupt changes: theory and application. Prentice-Hall, Upper Saddle River

9. Boracchi G, Roveri M (2014) Exploiting self-similarity for change detection. In: Proceedings of the international joint conference on neural networks, pp 3339–3346

10. Casillas MV, Garza-Castañón LE, Puig V (2012) Extended-horizon analysis of pressure sensitivities for leak detection in water distribution networks. In: 8th IFAC symposium on fault detection, supervision and safety of technical processes. Elsevier, pp 570–575

11. Covas DIC, Ramos HM (2010) Case studies of leak detection and location in water pipe systems by inverse transient analysis. J Water Resour Plan Manag 136(2):248–257

12. Cugueró-Escofet MÀ, Puig V, Quevedo J (2017) Optimal pressure sensor placement and assessment for leak location using a relaxed isolation index: application to the Barcelona water network. Control Eng Pract 63:1–12

13. Cugueró-Escofet P, Blesa J, Pérez R, Cugueró-Escofet MA, Sanz G (2015) Assessment of a leak localization algorithm in water networks under demand uncertainty. IFAC-Pap 48(21):226–231

14. Djelloul I, Sari Z, Latreche K (2018) Uncertain fault diagnosis problem using neuro-fuzzy approach and probabilistic model for manufacturing systems. Appl Intell 48(9):3143–3160

15. Eliades DG, Polycarpou MM (2012) Leakage fault detection in district metered areas of water distribution systems. J Hydroinf 14(4):992–1002

16. Fagiani M, Squartini S, Gabrielli L, Spinsante S, Piazza F (2015) A review of datasets and load forecasting techniques for smart natural gas and water grids: analysis and experiments. Neurocomputing 170:448–465

17. Ferrante M, Brunone B, Meniconi S (2007) Wavelets for the analysis of transient pressure signals for leak detection. J Hydraul Eng 133(11):1274–1282

18. Goldenshluger A, Nemirovski A (1997) On spatial adaptive estimation of nonparametric regression. Math Methods Stat 6:135–170

19. Islam MS, Sadiq R, Rodriguez MJ, Francisque A, Najjaran H, Hoorfar M (2011) Leakage detection and location in water distribution systems using a fuzzy-based methodology. Urban Water J 8(6):351–365

20. Jung D, Lansey K (2014) Water Distribution System burst detection using a nonlinear Kalman. J Water Resour Plan Manag 141(5):04014070

21. Kallesøe CS, Jensen TN (2018) On the relation between leakage location and network pressures. In: 2018 IEEE conference on control technology and applications (CCTA). IEEE, pp 571–576

22. Kang J, Park Y, Lee J, Wang S, Eom D (2018) Novel leakage detection by ensemble CNN-SVM and graph-based localization in water distribution systems. IEEE Trans Ind Electron 65(5):4279–4289

23. Lambert A (2001) What do we know about pressure: leakage relationships in distribution systems? In: IWA conference system approach to leakage control and water distribution system management. Brno, Czech Republic

24. Lambert MF, Simpson AR, Vítkovský JP, Wang XJ, Lee PJ (2003) A review of leading-edge leak detection techniques for water distribution systems. In: 20th AWA convention, Perth, Australia

25. Li R, Huang H, Xin K, Tao T (2015) A review of methods for burst/leakage detection and location in water distribution systems. Water Sci Technol Water Supply 15(3):429–441

26. Luciani C, Casellato F, Alvisi S, Franchini M (2018) From water consumption smart metering to leakage characterization at district and user level: the gst4water project. In: Multidisciplinary digital publishing institute proceedings, vol 2, p 675

27. Martini A, Troncossi M, Rivola A (2015) Automatic leak detection in buried plastic pipes of water supply networks by means of vibration measurements. Shock Vib 2015:11–15

28. Mounce SR, Mounce RB, Boxall JB (2011) Novelty detection for time series data analysis in water distribution systems using Support Vector Machines. J Hydroinf 13:672

29. Page ES (1954) Continuous inspection schemes. Biometrika 41(1/2):100–115

30. Palau CV, Arregui FJ, Carlos M (2012) Burst detection in water networks using PCA. J Water Resour Plan Manag 138(February):47–54

31. Pérez R, Puig V, Pascual J, Peralta A, Landeros E, Jordanas L (2009) Pressure sensor distribution for leak detection in Barcelona water distribution network. Water Sci Technol Water Supply 9(6):715–721

32. 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 Eng Pract 19(10):1157–1167

33. Pudar RS, Liggett JA (1992) Leaks in pipe networks. J Hydraul Eng 118(7):1031–1046

34. Puig V, Ocampo-Martínez C, Pérez R, Cembrano G, Quevedo J, Escobet T (2017) Real-time monitoring and operational control of drinking-water systems. Springer, New York

35. Puust R, Kapelan Z, Savić DA, Koppel T (2010) A review of methods for leakage management in pipe networks. Urban Water J 7(1):25–45

36. Quiñones-Grueiro M, Bernal-de Lázaro JM, Verde C, Prieto-Moreno A, Llanes-Santiago O (2018) Comparison of classifiers for leak location in water distribution networks. IFAC-Pap 51(24):407–413

37. Ragot J, Maquin D (2006) Fault measurement detection in an urban water supply network. J Process Control 16(9):887–902

38. Rajeswaran A, Narasimhan S, Narasimhan S (2018) A graph partitioning algorithm for leak detection in water distribution networks. Comput Chem Eng 108:11–23

39. Romano M, Kapelan Z, Savić DA (2012) Automated detection of pipe bursts and other events in water distribution systems. J Water Resour Plan Manag 140(4):457–467

40. Ross GJ, Tasoulis DK, Adams NM (2011) Nonparametric monitoring of data streams for changes in location and scale. Technometrics 53(4):379–389

41. Rossman L (2000) Epanet 2 user's manual. United States Environmental Protection Agency

42. Sammut C, Webb GI (2017) Encyclopedia of machine learning and data mining. Springer, New York

43. Sattar AM, Ertuğrul ÖF, Gharabaghi B, McBean EA, Cao J (2019) Extreme learning machine model for water network management. Neural Comput Appl 31(1):157–169

44. Shekofteh M, Ghazizadeh MJ, Yazdi J (2020) A methodology for leak detection in water distribution networks using graph theory and artificial neural network. Urban Water J 17(6):525–533

45. Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22(8):888–905

46. Soldevila A, Blesa J, Jensen TN, Tornil-Sin S, Fernández-Cantí RM, Puig V (2020) Leak localization method for water-distribution networks using a data-driven model and Dempster–Shafer reasoning. IEEE Trans Control Syst Technol

47. Soldevila A, Blesa J, Tornil-Sin S, Duviella E, Fernandez-Canti R, Puig V (2016) Leak localization in water distribution networks using a mixed model-based/data-driven approach. Control Eng Pract 55:162–173

48. Soldevila A, Fernandez-Canti RM, Blesa J, Tornil-Sin S, Puig V (2016) Leak localization in water distribution networks using model-based Bayesian reasoning. In: European control conference (ECC). IEEE, pp 1758–1763

49. Soldevila A, Fernandez-Canti RM, Blesa J, Tornil-Sin S, Puig V (2017) Leak localization in water distribution networks using Bayesian classifiers. J Process Control 55:1–9

50. Stehman SV (1997) Selecting and interpreting measures of thematic classification accuracy. Remote Sens Environ 62(1):77–89

51. Tharwat A (2018) Classification assessment methods. Appl Comput Inform

52. Theodoridis S, Koutroumbas K (2009) Pattern recognition. Elsevier, Amsterdam

53. Thornton J, Lambert A (2005) Progress in practical prediction of pressure: leakage, burst frequency and consumption relationships. In: Conference proceedings. Halifax, Canada, pp 12–14

54. Vrachimis SG, Eliades DG, Polycarpou MM (2018) Leak detection in water distribution systems using hydraulic interval state estimation. In: 2018 IEEE conference on control technology and applications (CCTA). IEEE, pp 565–570

55. Wachla D, Przystalka P, Moczulski W (2015) A method of leakage location in water distribution networks using artificial neuro-fuzzy system. IFAC-Pap 48(21):1216–1223

56. Wilcoxon F (1945) Individual comparisons by ranking methods. Biometr Bull 1(6):80–83

57. Woźniak M, Zielonka A, Sikora A, Piran MJ, Alamri A (2020) 6G-enabled IoT home environment control using fuzzy rules. IEEE Internet Things J

58. Wu Y, Liu S (2017) A review of data-driven approaches for burst detection in water distribution systems. Urban Water J 14(9):972–983

59. Wu ZY, Sage P (2006) Water loss detection via genetic algorithm optimization-based model calibration. In: ASCE 8th annual international, vol 5, pp 1–11

60. Ye G, Fenner RA (2011) Kalman filtering of hydraulic measurements for burst detection in water distribution systems. J Pipeline Syst Eng Pract 2(1):14–22