Literature Review of Data Analytics for Leak Detection in Water Distribution Networks: A Focus on Pressure and Flow Smart Sensors

Xi Wan[1], Parisa Khorsandi Kuhanestani[2], Raziyeh Farmani[3], Edward Keedwell[4]

[1]Ph.D. Student, Centre for Water Systems, College of Engineering, Mathematics and Physical Science, University of Exeter, Harrison Building, North Park Rd., Exeter, Devon, EX4 4QF, United Kingdom (corresponding author). Email: xw355@exeter.ac.uk

[2]Ph.D. Student, Department of Marine and Fluvial Systems, University of Twente, 7522 NB Enschede, Netherlands. Email: P.khorsandikuhanestani@utwente.nl

[3]Centre for Water Systems, College of Engineering, Mathematics and Physical Science, University of Exeter, Harrison Building, North Park Rd., Exeter, Devon, EX4 4QF, United Kingdom. Email: R.Farmani@exeter.ac.uk

[4]College of Engineering, Mathematics and Physical Science, University of Exeter, Exeter EX4 4QF, United Kingdom. Email: E.C.Keedwell@exeter.ac.uk

ABSTRACT

Leakage detection is one of the important aspects of water distribution management. Water companies are exploring alternative approaches to detect leaks in a timely manner with high accuracy to reduce water losses and minimise environmental and economic consequences. In this article, a literature review is presented to develop a step-by-step analytic framework for the leakage detection process based on flow and pressure data collected from water distribution networks. The main steps of the data analytic for leakage detection are: setting up the goals, data collection, preparing the gathered data, analysing the prepared data, and method evaluation. The issues of concern for each step of the proposed leakage detection framework are analysed and discussed. The smart sensor-based leakage detection methods can be categorised as data-driven methods and model-based methods. Data-driven methods can be further categorised as statistical process control-based methods, prediction-classification methods, and clustering methods. Hydraulic model-based methods can be further categorised as calibration-based methods, sensitivity analysis, and classifier-based methods. The advantages and disadvantages of each method are discussed, and suggestions for future research are provided. This review represents a new perspective on the subject from five aspects: 1) most of the leakage detection methods are focused on burst detection, and different types of leakages should be considered in future research; 2) it is important to consider data uncertainties, and more robust real-time leakage detection methods should be developed; 3) it is important to consider hydraulic model uncertainties; 4) unrealistic assumptions should be addressed in future research; 5) spatial relations between sensors could provide more information and should be considered.

INTRODUCTION

Leakage is the loss of water from the supply network through uncontrolled actions. In addition to water loss, there are other negative consequences caused by leakage (Colombo et al. 2009; Farah and Shahrour 2017), such as: 1) potential risks to public health (Romano et al. 2011) due to the entrance of contaminants from the environment into the pipes with negative pressure and changes in water quality (Xu et al. 2014), 2) environmental issues due to energy used by pumps to deliver water to compensate for the pressure drop, and chemicals used to treat the water in the treatment plant, and 3) leaked water that ends up in surface water with a potential negative impact on living organisms due to high chlorine concentration. All of these lead to great socio-economic losses (Colombo et al. 2009; Farah and Shahrour 2017; Romano et al. 2011; Wu et al. 2010a), and thus, timely detection and localisation of leakage events in Water Distribution Networks (WDNs) have received considerable attention for more than two decades (Zaman et al. 2020).

Leakage detection and localisation are the processes to identify leakage in the WDN and specify the leakage location. Timely detection of leaks could decrease the amount of water loss from the system. It can also have other benefits (Bohorquez et al. 2020; Puust et al. 2010). For example, it can (i) reduce environmental impacts by reducing water losses, (ii) allow planned interruption of supply and therefore reduce the impact on customers and (iii) reduce financial costs by reducing the level of pumping and financial losses linked to the amount of water lost. After detecting a leak, the location of the leakage should be determined, so that repair can take place. The determination of leakage location has three different phases. The first phase is to "localise", i.e., limit the location of a leak to a specific district.

The second phase is to "locate" the pipes in a district area where leaks occur. The third phase is to "pinpoint", i.e., limit the leakage location to an area with a small radius of 2-3 feet (Qahtani et al. 2020).

Almost all leakage detection methods can be broadly categorised into hardware-based methods and software-based methods (Ismail et al. 2019). Hardware-based methods, also called direct methods (Zaman et al. 2020) or passive methods (Chan et al. 2018), usually rely on hardware devices to detect leak events. Based on the principles that apply to the hardware devices, hardware-based methods can be further divided into acoustic techniques and non-acoustic techniques (see Fig. 1). The detailed information about the hardware-based methods can be found in Ismail et al. (2019) and Chan et al. (2018). Although the accuracy of the hardware-based methods is increasingly high, they are costly, time-consuming and labour-intensive, as expensive equipment and professional staff are needed. Furthermore, the results could be influenced by pipe materials (e.g. acoustic methods), soil types and conditions (e.g. infrared thermography, ground penetrating radar), and other factors depending on the equipment. Therefore, hardware-based methods are mostly used in the third phase of leakage detection and localisation, i.e. "pinpoint".

Software-based methods, also called indirect methods (Zaman et al. 2020) or active methods (Chan et al. 2018), could detect leaks by inference from internal pipeline parameters (such as pipe flow and pressure data) rather than detect leak-related information (e.g. leak noise, infrared radiation) directly. Software-based methods use computer software systems to monitor hydraulic parameters to detect possible leaks continuously. Unlike hardware-based methods that try to pinpoint the leakage location accurately, software-based methods aim to detect leakage and limit the area to a specific district. Based on the hydraulic state of the pipeline system, the software-based methods can be classified as transient-state methods and non-transient methods.

Colombo et al. (2009) and Abdulshaheed et al. (2017) provided a comprehensive review of the transient-state leakage detection method. Transient analysis is based on the idea that any changes (e.g. blockage, leakage) in the pipe's physical structure will alter the flow and pressure response. To adequately capture the transients at all time scales, the system requires many measurement points with high sampling frequency, which results in a costly, labour-intensive process. Moreover, this technique often relies on complex transient simulation models and is mainly applied to a single pipeline to predict its features (Bohorquez et al. 2020; Keramat et al. 2019). This type of analysis is computationally expensive and unsuitable for real-time monitoring of large urban areas.

Non-transient methods, which is the focus of the present study, can be further classified as hydraulic model-based methods and data-driven methods according to whether a hydraulic model is used. Compared with transient-state methods, non-transient methods could use monitoring data with a much lower sampling rate (e.g. 5 minutes, 15 minutes), which is readily available. Non-transient methods provide a promising solution for the long-term monitoring of large-scale WDNs. When a leak occurs, it will change the hydraulic behaviour in the WDS, and flow and pressure readings also will change. The flow will increase by additional demand, which results in larger head loss, and leads to different pressures within the network. Leakage detection is based on the difference between the predicted hydraulic parameter in the absence of leakage provided by hydraulic model or data-driven methods and the field observations collected by sensors.

Therefore, real-time monitoring of changes in water distribution networks to detect leaks is one of the most promising methods. The monitoring is often performed by installing pressure sensors in different parts of the system or installing flow sensors in the transmission mains. With the rapid development of the internet of things and big data technologies in recent years, smart and intelligent water systems could be more connected and operated with more data in real-time to achieve maximum efficiency and effectiveness. The sensors create big data, and by applying appropriate data analytics to them, valuable information could be obtained and make the detection and localisation of leakages possible. However, the focus on big data is relatively new in the water industry, and most of the decision making is done either independently or with limited use of available data. Therefore, this paper focuses on smart sensor-based leakage detection systems that use time-series data from pressure and flow sensors to detect leakage in distribution networks.

The literature review on existing methods allows researchers to form reasoned, logical and confirmed arguments (Denyer and Tranfield 2006). Several literature review articles have been published on leakage management with the main focus on detection. Puust et al. (2010) presented a review of leakage management methods and classified these methods into three groups: leakage assessment methods, leakage detection methods and leakage control models. The authors concluded that future works should focus on the real-time models for pipe networks. Gupta and Kulat (2018) carried out a similar review to Puust et al. (2010) and highlighted that more effort is needed in online monitoring and online leakage detection. El-Zahab et al. (2016) reviewed leakage detection methods focusing on: 1) classification of leak detection phases (identification, localisation, and pinpointing leaks), 2) sensor installation type (static and dynamic leak detection systems). Wu and Liu (2017) reviewed data-driven methods using data from Supervisory Control and Data Acquisition (SCADA) systems but only for burst detection. They categorised the methodologies into three groups: classification method, prediction method, and statistical method. Hu et al. (2021a) reviewed model-based and data-driven approaches for leakage detection and location from the aspect of methodology. Hu et al. (2021a) provide a good overview of methodologies that have been developed for leakage detection, but in this paper, a more comprehensive step-by-step analysis of the process for leakage detection is provided, including the data pre-processing techniques, the types of case study, the size of leaks that are possible to be detected, etc. In order to thoroughly discuss those topics, this paper provide a step-by-step analytic framework for real-time leakage detection process based on big data gathered from pressure and flow sensors, from the perspective of water resources planning and management.

The aim of this review is to clarify the state of knowledge, identify research gaps, and form a consensus on the subject. A five-step framework has been developed to analyse and compare a suite of leakage detection and localisation methods for low-frequency (compared with data for transient analysis) pressure and flow data measured by in situ sensors (Fig. 2). It is no doubt that goal identification is the prerequisite of data analysis. After installing sensors in the distribution network, the first step is to collect data for the following analysis. Identifying the characteristics of the collected data is a crucial stage for data analytics-based studies, and suitable data analytic methods could be chosen accordingly. By applying the appropriate methodology, the information will gain more value and be used in decision making. Based on the data collected from the SCADA system and the analysis of the characteristics of collected data, different methods have been applied to detect and localise leakage events. Each step will be analysed step by step in the following research.

GOAL IDENTIFICATION (STEP 0)

Identifying the needs and goals of the end-user is the first step in the leakage detection and localisation framework. Setting goals has a direct impact on the selection of the most suitable detection methods. In most proposed methods, leakage identification and localisation are the two goals of leakage detection based on data analytics. Tables 1 and 2 provide the summary of steps 0 and 1 of data-driven and hydraulic model based leakage detection methods, respectively. Data-driven methods model the system behaviour based on historical data mining, and model-based methods use a well-calibrated hydraulic model to represent the current state of a network. Therefore, for leakage identification, hydraulic model-based methods are preferred when the amount of historical data is limited, and data-driven methods could be more efficient and provide more accurate predictions when a long-term monitoring dataset is available. For leakage localisation, model-based methods are more preferred since the hydraulic model can provide more topological information of the network.

Leakage identification usually ends with binary results (i.e. alarm on or off) that represent whether or not a leakage is happening in the system. In order to provide more information for the operator, Mounce et al. (2007, 2010) have provided fuzzy values and probability values between 0 to 1 to represent the likelihood of a leakage event. Ye and Fenner (2011, 2014) defined the burst size as the difference between the predicted value and the observed value. The determined burst size may not be very accurate due to prediction errors, measurement errors, unpredictable random consumers' behaviour, etc. The results of leakage localisation are presented as a leak map to compare the predicted leak area and the actual leak location. Visualisation of the results helps the decision-makers to gain knowledge and make informed decisions.

DATASET GENERATION (STEP 1)

One of the essential stages of data analytics-based studies is data collection and data characteristics. Data specifications should be based on the needs of the study and the goals that have been set. Most of the studies have applied their method to a sample case to validate or explain the presented methodology of leakage detection. Those algorithms are usually tested on three kinds of datasets:

- A synthetic dataset generated by a hydraulic model (such as EPANET, WNTR etc.) that usually made simplifications to the network's condition and customers' behaviour (such as following a very regular pattern or assuming that consumers' behaviour are known in advance);
- Engineered tests operated in real water distribution systems by opening fire hydrants that usually simulated burst events with sudden water loss and short duration.
- Historical monitoring data that contain real leakage events.

The artificially synthesised dataset allows researchers more flexibility to adjust parameters and model different types of leakage events. However, it is difficult to model the high uncertainties in a real-life network. The simplifications of simulations can generate a clean and easy-to-learn dataset that is beneficial for evaluating the tested algorithms and leaves a question of the suitability of the algorithms for real-life networks. Therefore, these methods should be validated in engineered test datasets or historical datasets. As shown in Table 1, engineered tests that can introduce artificial leakage to the real system are widely-used when researchers evaluate their methods. It should be noted that the engineered tests used in the current studies only simulated burst events but not incipient leakage. However, incipient leakage can cause more water loss than bursts due to a longer awareness time. Furthermore, incipient leakages may last for days, or even weeks, unlike burst events that lasted for a few hours simulated in the engineered tests. Therefore, it is crucial to generate an early warning for incipient leakage before it reaches its maximum level. Historical monitoring datasets that contain real leakage events can reflect the efficacy of detection methods in real-life scenarios. However, it is hard to know the exact starting time of leakages in real life, and the information can only be inferred from the maintenance work or customer contacts, which brings difficulty for method evaluation. Within this context, the Battle of the Leakage Detection and Isolation Methods (BattleDIM) (Vrachminis et al. 2020) provides a hydraulic model called L-Town that contains two years of real-life demand data. The datasets of BattleDIM provide two years of monitoring dataset, including flow data, pressure data, and demand data. Different types of leakages are modelled in the system, including background leakages, gradual leakages, and bursts. These datasets could be a good option to evaluate the performance of leakage detection and localisation methods.

The commonly used data are flow and pressure monitoring data, which are critical hydraulic parameters that change with any alteration in the distribution system. In some cases, the flow values are reported as the average value, while pressure values are reported instantaneously (Mounce et al. 2012). Therefore, the data shows smoothed flow values and missing some changes in flow between time intervals (Farley et al. 2013). Compared with flow time series, pressure time series have more variation in their profile than flow time series (Romano et al. 2011). Furthermore, flow instruments are usually installed at inlets, and they are sensitive to downstream changes, while pressure values are affected by head loss and pressure changes upstream and around the instrument position (Geiger 2005). Therefore, in the experiment conducted by Ye and Fenner (2011), pressure-based detection seems less sensitive to a burst event than flow-based detection, especially when a pressure sensor is remote from the burst location. It can also be observed from Tables 1 and 2 that most of the burst identification methods preferred to use flow data. However, the installation of flow sensors is more expensive than pressure sensors (Romano et al. 2011). Hence, usually, there are fewer flow data available than pressure data. It has been suggested that the pressure data can be used to provide additional information as a way of confirming the flow-based detection results (Ye and Fenner 2011). Therefore, pressure data are preferred when the goal is leakage localisation.

DATA PREPARATION (STEP 2)

Raw monitoring data may contain lots of noises, missing data, or data from faulty loggers. Therefore, it is necessary to perform data pre-processing before data is analysed by data analytics. Furthermore, it is difficult to construct a proper function to model the pattern of the raw monitoring data because it is highly nonlinear. In addition, the variation over different weekdays makes the flow and the pressure pattern more complicated to model. While some papers have tried to model water usage patterns directly, most papers adopted pre-processing techniques before leakage detection. Moreover, different methods have different requirements that need to be prepared before analysing. Tables 3 and 4 summarise steps 2-4 of each leakage detection method.

Data-Driven Methods

As shown in Table 3, the most frequently used data preparation procedures for data-driven methods are:

1. Data correction: Data collected by the sensors is a type of big data, and due to its real-time nature, sometimes incomplete/incorrect data may exist in the time series, due to missing data, data from faulty loggers, erroneous timestamps etc. To ensure a continuous data stream, in some studies, the missing data are replaced by an alternative value that is calculated through a statistical process such as a filter interpolation (Mounce et al. 2002; Mounce and Machell 2006; Romano et al. 2014). Furthermore, statistical tests could be applied to the time series to ensure that an adequate amount of good quality data is available for the analysis (Romano et al. 2014).
2. Data de-noising: Recorded pressure and flow data are usually accompanied by noise. The presence of this noise may cause some small leaks to be undetected or cause false alarms. Therefore, in some studies, this noise is removed in the pre-processing stage. Misiunas (2006) used an adaptive recursive least squares filter, and Romano et al. (2014) used discrete wavelet transform to remove noise from data.
3. Data selection: In most studies, a range of normal data (data without leakage events) is needed to train machine learning models or used as a reference library so that the normal behaviour of the distribution network can be accurately represented. However, data collected from real WDNs usually contains both leak and non-leak events. Therefore, data selection is needed to ensure the performance of the detection method. For example, Palau et al. (2012) used an iterative procedure to eliminate outliers during Principle Component Analysis (PCA) model construction. Wu et al. (2020) used an abnormal subsequence searching (ASS) algorithm to search and remove the abnormal subsequences in the library.
4. Data reformatting: Different techniques require different formats for the data, such as normalisation, label information assigning, time-series restructure etc. When the analytic method is sensitive to the numerical ranges of the variables, such as PCA (Palau et al. 2012), mean centring and scaling processes are needed to normalise the data to ensure that the data falls within the same range. Label information such as time of day, day of the week could provide more information to the machine learning model and improve the prediction accuracy. Furthermore, given a sequence of numbers for a time series dataset, data needs to be pre-processed (e.g. a tapped delay line format) to prepare for neural network presentation.

Hydraulic Model-Based Methods

As shown in Table 4, the most frequently used data preparation procedures for hydraulic model-based methods are:

1. Model calibration: For hydraulic model-based methods, hydraulic model calibration is unavoidable for hydraulic model-based methods. The aim of the model calibration is to develop the best values for the unknown model parameters, so that the hydraulic model could reasonably represent the performance of the WDN. A poorly calibrated model could result in significant errors in leak detection. A reliable hydraulic model requires structural and hydraulic data for calibration and validation (Giorgio Bort et al. 2014). Due to the fact that it is extremely difficult to accurately obtain the roughness and diameter of every pipe and water demand at each node, these parameters should be properly calibrated and validated before the hydraulic model can be used. Evolutionary methods and least-squares are the most frequently used methods for model calibration (Sanz et al. 2016).
2. Data generation: With the availability of a well-calibrated hydraulic model, various leakage scenarios (including the non-leakage scenario) under different boundary conditions can be modelled (Soldevila et al. 2019). By simulate different leakages with different locations and different scales, a hydraulic model could provide numerous training example for leakage area classification. The leakage can be modelled as an emitter flow that represented as a function of the pressure at the junction node, given as

$$Q_i(t) = k_i[P_i(t)^{\alpha}] \quad (1)$$

where Q_i is the leak flow at node *i* at time *t*, k_i is the emitter coefficient at node *i*, $P_i(t)$ is the nodal pressure at node *i* at time *t*, α represent the emitter pressure exponent.

Historical leak-free monitoring measurements (such as pressures, flows, reservoir conditions, etc.) must be provided to the hydraulic model as boundary conditions. The network behaviour can be described by steady-state models concatenated in an extended period simulation (EPS) (Perez et al. 2014). In addition, pressure-driven analysis (PDA) (Wagner et al. 1988) in EPANET provides a more realistic representation of the pressure-leakage relationship.

3. Zone division: Usually, a small number of monitoring devices are equipped in WDNs. Thus, it is hard to locate the exact location of leakage with limited monitoring devices since leakage that happened at neighbouring pipes might have a very similar influence on available devices. Therefore, in order to locate the leakage in a small possible area within a district meter area (DMA), one solution is to divide the DMA into small zones. Zhang et al. (2016) used k-means clustering algorithm to divide WDNs into k leakage zones. Wu et al. (2022) used fuzzy c-means (FCM) to cluster pipes and place the sensors. Romero et al. (2022) adopted a method called graph agglomerative clustering (GAC) to cluster the network based on its topology.

DATA ANALYTIC METHODS (STEP 3)

Various techniques have been explored to mining the monitoring data and to provide effective solutions for leakage detection. As mentioned before, leakage detection methods can be categorised as data-driven methods and hydraulic model-based methods.

Data-Driven Methods

Currently, data-driven methods are mainly used for leakage identification (especially for burst detection) instead of leakage localisation. With a large amount of historical data, the pattern can be analysed by statistical methods or learnt automatically by machine learning models. If the characteristics of new data are substantially different from historical data, it can be inferred that an abnormal event occurred in the distribution system. For example, a sudden pressure drop and flow increase are the most frequently used criteria for burst events. Fig. 3 shows the flowchart of data analytic steps for data-driven leakage detection methods, and Table 3 summarises the techniques used in each method. Based on the principle of the data analysis techniques, data-driven methods can be further categorised into three categories: statistical process control (SPC)-based methods, prediction-classification methods, and clustering-based methods.

SPC-based methods

Statistical process monitoring charts, also called control charts, with a set of control limits, are used to display and detect the unusual variability in the data. SPC methods are the most intuitive and simple but powerful methods used to monitor the unusual behaviour of a process. These charts contain three characteristics: a target representing the mean value for the in-control process, upper control limit, and lower control limit used to determine the in-control limits. The control limits can be set by calculating the statistical characteristics of the historical data. Data that is outside these thresholds are assumed as invalid or abnormal. For example, the well-known "3-sigma" method is a Shewhart-type method, which means that the data three standard deviations from the mean are considered under normal conditions.

The most common SPC methods include univariate methods such as Shewhart chart (Loureiro et al. 2016), Western Electrical Company (WEC) rules (Ahn and Jung 2019; Jung et al. 2015), cumulative sum (CUSUM) control chart (Misiunas et al. 2006), exponentially weighted moving average (EWMA) (Jung et al. 2015), and multivariate methods that consider the correlation between data from multiple sensors, such as Hotelling T^2 (Palau et al. 2012), multivariate EWMA (Jung et al. 2015), and multivariate CUSUM (Jung et al. 2015). Shewhart-type approaches provide effective detection of large faults, while CUSUM and EWMA are more sensitive in detecting small changes but do not guarantee to detect large faults (Harrou et al. 2020). Some of the SPC methods can be viewed as a simpler version of prediction model-based methods discussed in the next category. For example, Shewhart uses the mean of historical value as the predicted value for the next data point, and EWMA uses an exponentially weighted moving average value.

It should be noted that some assumptions that underlie the quality control process are: 1. Data comes from a single statistical distribution; 2. The data distribution is a normal (Gaussian) distribution; 3. The errors are uncorrelated over time. It is clear that none of these assumptions holds true in the raw monitoring data. Therefore, some researchers (Jung et al. 2015; Loureiro et al. 2016) assumed that the data at the same time every day comes from the same distribution and reformat the data before applying the SPC methods. The Minimum Night Flow (MNF) analysis proposed by Farah and Shahrour (2017) also implies this idea, in which the minimum value of a day is checked based on the calculation of moving average and moving standard deviation. Loureiro et al. (2016) improved SPC methods by using a quantile-based approach instead of a sample mean to consider the asymmetric behaviour of flow data. Besides, instead of considering each data point separately, Palau et al. (2012) divided a day of monitoring data into different time periods (such as morning, afternoon, and night) and used PCA to compress the data and to reduce the unnecessary variability. Then, Hotelling T^2 and distance to model can be calculated to determine the outliers. However, this method takes a relatively long time to make the decision, which may not be conceived as very effective for burst detection.

Prediction-classification methods

Prediction-classification methods are the most common approaches in the literature. While SPC-based methods try to construct boundaries directly for the monitoring data points based on unrealistic assumptions, prediction-classification methods construct more complicated and accurate models that can represent the expected behaviour of the distribution system in a healthy state. The prediction model is trained with the collected historical data and then gives predicted values. If there is no leakage in the future, there will be a reasonable match between predicted and measure values. Therefore, anomalies representing leakage events can be detected by analysing the residuals between the observed value and the predicted value (see Fig. 3).

Various approaches have been explored to detect leakage events in WDNs. Ye and Fenner (2014) have proposed the weighted least squares with the expectation-maximisation algorithm for burst detection. They (Ye and Fenner 2011) also explored the application of Kalman Filter (KF) for burst detection and achieved very good results. However, one shortcoming of these two approaches is that they considered data within the same day separately. If the sampling rate is 15 minutes, it needs to build 672 independent models to eliminate diurnal patterns. If complex patterns are considered, such as weekly patterns, more models need to be built. Thus, building a large number of filters cannot radically solve the problem of equal-state assumption.

Machine learning techniques can learn from data without relying on rules-based programming. Mounce et al. (2002, 2010, 2011) have used support vector regression (SVR) and artificial neural network (ANN) for leakage detection on real data from a water distribution system. Romano et al. (2014) provided an online system for leakage detection using ANN combined with the Bayesian inference system (BIS). However, a standard ANN does not share features across different steps of time series. In contrast, the recurrent neural network (RNN) is widely recognised as a suitable method to deal with sequential data due to its ability to connect previous information to the present task. Therefore, Wang et al. (2020) used the long short-term memory (LSTM) network - a special kind of RNN - for flow prediction and the detection of burst events.

Besides the methods that have been explored in water leakage detection (see Table 3), various kinds of regression or prediction models can be applied in the prediction stage (Han et al. 2019), as long as the models are capable of time series modelling. Prediction-classification detection methods highly rely on the accuracy of the prediction model (Wu and Liu 2017). However, time-series predictors generally have no inbuilt mechanism for subsequent classification. Thus, additional classification methods, such as control charts, have been used for the final alarm raising or decision making. However, setting the threshold for event detection is not a trivial issue. The determination of thresholds often depends on experience, which greatly influences the detection effect (Wang et al. 2020). Additionally, most of these prediction-classification methods need to be updated regularly to adapt to time changes (Mounce et al. 2010).

Clustering-based methods

Clustering-based methods are based on comparing time series subsequences or their representations, using a reference of normality, without the need for fitting a prediction model. Clustering analysis is used to create clusters by grouping points or subsequences that are similar to each other and separating dissimilar points or subsequences into different clusters. Abnormal subsequences are those that are dissimilar to normal subsequences, and they can be determined based on the distance to the centroid of the cluster of normal sequences belongs.

Wu et al. (2016) used cosine distance to calculate the dissimilarity between vectors combined with the information from different sensors. Aksela et al. (2009) proposed a method based on the self-organising map (SOM) to detect leakage by finding similarities between flow data from other weeks, facilitated by a leak function that describes the relationship between the confidence in the existence of a leak and the distance between flow meters and leakage locations. Wu et al. (2020) proposed a shape similarity-based (SSB) method that detected bursts by analysing the shape of flow time series data within the same period from different days. Huang et al. (2018) applied dynamic time warping (DTW) to study the similarity of daily water demand and found the most unusual daily pattern.

The construction of the reference library is the most critical step of clustering-based methods. In order to consider various uncertainties in the data, such as weekday patterns, weekend patterns, holiday patterns, etc., a large amount of historical data may be needed. Furthermore, those discord discovery techniques (e.g. Wu et al. (2020), Huang et al. (2018)) require the users to specify the length of the leakage event in advance, which in many cases may not be known and could only be determined by experience. It should be noted that most of the methods are designed for burst detection. Most of the clustering-based methods are based on the assumption that leakage events will cause unusual shapes in the data. However, the shape of the pattern of incipient leakage may stay the same during the beginning stage but grow in the long term. Thus, the ability of clustering methods for incipient leakage detection is still a question.

Hydraulic Model-Based Methods

Hydraulic model-based approaches rely on a hydraulic model of a network. The accuracy of these models depends on their calibration, and a prerequisite of accurate leakage localisation is a well-calibrated hydraulic model. By comparing the simulations generated by the well-calibrated hydraulic model and the data collected from pressure and flow sensors, the leakage can be detected, and the most probable area of the network can be found. Currently, hydraulic model-based methods are mainly used for leakage localisation. In most studies, the leakage detection is based on the results that a leakage event has already been known to exist in the network. Based on the principles used to detect leakage, hydraulic model-based methods can be further categorised as calibration-based methods, sensitivity analysis-based methods, and classification-based methods. Details of these methods can be found in Table 4. The general steps of hydraulic model-based leakage detection methods have been described in Fig. 4.

Calibration-based methods

Leakage detection based on model calibration is defined as an inverse problem of parameter identification of the hydraulic model. The leakage detection is initiated by obtaining the well-calibrated hydraulic model and the field measurements. Model calibration is used to minimise the discrepancies between the observed flow and pressure values and the values simulated at junctions in the hydraulic model affected by possible leaks. After selecting the optimisation criteria and the optimisation objective, an optimiser will be used to seek the best solution from all possible solutions automatically. Possible solutions are usually represented as a number of leakage nodes with positive emitter coefficients (Wu et al. 2010a). Table 4 summarises the objective function and the optimisation method that have been explored in each study.

Misiunas (2006) searched for all the locations, and the node with the smallest objective value (i.e. the sum of difference squares) is declared to be the burst position. This method is based on the leakage identification procedure, the amount of water loss needs to be estimated, and the demand value is assigned uniformly to all nodes. However, the leakage demands should be pressure-dependent, which is more in line with reality. Therefore, Wu et al. (2010b) developed a pressure-dependent leakage detection (PDLD) method that uses pressure-dependent emitter flow at a junction to represent a leakage event. A Genetic algorithm (GA) was then used to search for the optimal solution. Then, Wu et al. (2010b) provided an application report of the PDLD method to two water systems and proved its effectiveness compared with acoustic leak loggers.

The number of decision variables is directly related to the number of candidate leak locations in a DMA, which means that the optimisation algorithm needs to solve a nonlinear inverse problem with thousands of decision variables for a medium-sized system. When there are a large number of decision variables because multiple combinations of decision variables may generate equally fit solutions and result in inaccurate localisation results (Sophocleous 2019), which greatly limits the applicability of calibration-based methods. Although the decision variables have been reduced by specifying the maximum number of possible leaks within a system, this information heavily relies on engineering judgment. Therefore, Sophocleous et al. (2018, 2019) introduced a search space reduction stage before leak localisation to reduce the search area of the optimisation. A real case from the UK is investigated by this method and proved its effectiveness. However, these methods have only been demonstrated in single-leak cases because of the combinatorial complexity and a large number of decision variables. In order to take into account multiple leakage scenarios, Berglund et al. (2017) proved that multiple leaks could be modelled as a linear combination of single-leak scenarios under certain limitations (e.g. leak coefficients, leak number). Based on linear programming (LP) and mixed-integer linear programming (MILP), a linear combination of leaks can be determined to approximate the observed pressure values.

The advantages of calibration-based methods are: 1) the leak position can be accurately determined; 2) the leak amount can also be determined. However, one major drawback of these methods is that they are computationally demanding and have limited real-time applicability (Berglund et al. 2017). Sanz et al. (2016) proposed an online leakage detection method based on model calibration. The leak can be determined based on the demand difference between the hydraulic model being calibrated before and after. However, this method is computationally demanding. The calibration-based methods are mostly applied to small networks since the effectivity of the optimisation algorithm can be greatly limited.

Sensitivity analysis-based methods

Sensitivity analysis-based methods detect leaks based on the comparison of modelled data versus observed data. The quantification of the difference between the actual pressure measurements with the predictions predicted by the hydraulic model is called pressure residual. After the model has been calibrated using the historical data, the pressure residual set for n_s sensors can be obtained by calculating the theoretical pressure difference between the non-leak scenario, $\hat{p}_0 \in \mathbb{R}^{n_s}$, and all potential leak scenarios, $p \in \mathbb{R}^{n_s}$, simulated at each junction in turn:

$$r = p - \hat{p}_0 \tag{2}$$

Based on evaluating the theoretical effect of all potential leaks f_i of all monitored nodes, p_i, the sensitivity matrix S can be determined as:

$$\mathrm{S} = \begin{bmatrix} \frac{\partial p_1}{\partial f_1} & \cdots & \frac{\partial p_1}{\partial f_{n_p}} \\ \vdots & \ddots & \vdots \\ \frac{\partial p_{n_s}}{\partial f_1} & \cdots & \frac{\partial p_{n_s}}{\partial f_{n_p}} \end{bmatrix} \tag{3}$$

where n_s is the number of sensors, n_p is the number of potential leaks (network nodes). However, it is extremely difficult to calculate the sensitivity matrix S analytically for a real network since the WDS is a nonlinear system without an explicit solution (Perez et al. 2014). Therefore, many ways have been proposed to approximate the sensitivity matrix. For example, Giorgio Bort et al. (2014) and Okeya et al. (2015) estimated the burst flow before leakage localisation, and the estimated burst flow was simulated in turn at each node to obtain the sensitivity matrix. Perez et al. (2011, 2014) introduced the same leakage in each node and recorded the pressure increment to approximate the sensitivity matrix. To construct a more robust sensitivity matrix, Farley et al. (2013) used the chi-squared value of pressure increment. Furthermore, to consider the uncertainties that exist in real-life, Pérez et al. (2011) applied a threshold to the sensitivity matrix so that only the strong relations between leaks and pressure sensors could be considered.

The possible leak area can be determined by ranking the sensitivity of sensors to leaks and compared with the observed pressure residual. Giorgio Bort et al. (2014) performed PCA analysis on the sensitivity matrix to rank the measurement nodes according to the most important feature. Perez et al. (2014) and Steffelbauer et al. (2022) located the most probable leak nodes by identifying the largest correlation values between the observed pressure residual and sensitivity matrix. Theoretically, the residuals should be zero under non-leakage scenarios. However, due to the existence of measurement errors, calibration errors, random customer behaviour etc., the residual will not remain zero even under healthy conditions. Furthermore, the sensitivity of pressure sensors to different leak scenarios is hard to quantify using a constant value, which reduces the accuracy of leak localisation.

Classification-based methods

With the development of machine learning techniques, the classification of leakage scenarios can automatically be trained by a classifier. In the first stage, the pressure map or pressure residual map of each leakage scenario can be generated by the hydraulic model and used as training data. In the second stage, the training data will be processed and fed to train a classifier. After a burst is detected, observed pressure values will be processed, and the trained classifier could be used to determine the leak area.

Zhou et al. (2019a) used the fully-linear DenseNet (FL-DenseNet) to extract features in pressure patterns for burst localisation. Javadiha et al. (2019) used a convolutional neural network (CNN) to learn the different pressure residual maps. Since the number of pressure sensors in a system is limited, some node leaks may present a similar leak signature and can be indistinguishable. Therefore, Soldevila et al. (2016) used a node grouping procedure prior to the k-Nearest Neighbour (kNN) classifier training. Zhang et al. (2016) used K-means clustering to divide the network into different k zones based on the pressure residual matrix generated by the hydraulic model. Then, the leakage events were represented by adding a random leakage demand to the junction selected by the Monte-Carlo method in the hydraulic model. Training samples are generated by the hydraulic model and used to train the M-SVM model. Romero et al. (2022) used an image coding procedure called Gramian angular field (GAF) to transform pressure vectors into images, and the task of leakage localisation has been transformed into image classification. A set of deep neural networks (DNNs) are organised hierarchically to obtain a classification tree to localise the leakage area. Zhang et al. (2022) used FCM to divide network into different zones and combined the extreme gradient boosting (XGBoost) to identify the leakage zone. The results showed superior performance than the back-propagation neural network (BPNN). More information can be found in Table 4.

With the benefit of a hydraulic model, a large number of training data can be generated and provided for classifier training. Once the classifier is well-trained, the results can be generated efficiently. However, a well-calibrated hydraulic model is hard to be maintained to reflect the real-time condition of the network. Any changes in the network will cause inaccurate estimation, such as the addition or elimination of any element (pipes, valves, tanks, etc.), the changes of pipe roughness coefficient and the changes of pipe diameter caused by increasing pipe ages. Furthermore, consumers' demand is hard to determine and difficult to adjust its real-time variation. Currently, hydraulic model-based methods have not reached the maturity of real-time monitoring for WDNs.

PERFORMANCE EVALUATION (STEP 4)

Method evaluation has a critical role in method development, and different goals need different metrics. Each method should be evaluated before being applied to real life. In this section, model validation and method performance with real-life data will be analysed.

Leakage Identification

For leakage identification, the process usually ends up with binary classification. Essentially, each data point or a data subsequence needs to be labelled as an anomaly or not. If leakage happens and data points during that time period are identified as anomalies, this case is a true positive. If there is no leakage happening and the detection results showed negative all the time (meaning the system is healthy), then it's called a true negative. However, there are cases the detection method can fail. If the system is healthy, but an alarm is rising, this case is a false positive. If leakage happens but the detection results show negative, this case is a false negative. Therefore, as one of the most comprehensive ways, confusion matrices (Alla and Adari 2019) have been widely used to evaluate leakage identification methods' performances. True Positive Rate (TPR) and False Positive Rate (FPR) are the two most commonly used criteria for the evaluation of leakage detection methods. This is because failure events rarely occur in real-life scenarios, resulting in considerable parts of the observed data being labelled as normal, and a few parts of it are labelled as abnormal. As such, other indicators such as F1 score, Receiver Operating Curve (ROC) or area under the curve (AUC) can be used to evaluate such biases in data. For prediction-classification methods, in the prediction stage, additional evaluation metrics such as rooted mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) should be used to measure and quantify the prediction error. It is important to emphasize that the prediction error will never be zero because no model can perfectly predict the future. From Table 5, it could be observed that the current methodologies have achieved very high accuracy, but it should be noted that most of them focused on burst events. Furthermore, only a few of them are applied in real life.

Furthermore, the false alarm presents a serious issue. Some results could have false positives every day (Jung et al. 2015, Jung and Lansey 2015, Xu et al. 2020), it is clearly impossible to raise alarms at every false positive point. It is well-known that one abnormal data point cannot solely represent the leakage event, because it has a high probability that it is generated by the data noise or the random behaviour of water consumption, while continuous disruptive data is more suitable to indicate the occurrence of a leakage event. For example, Mounce et al. (2007) combined ANN with Fuzzy Inference System (FIS). Mounce et al. (2011) applied a time window for the detection results, and an alarm will be raised only if enough anomalies occur within a moving event window. Romano et al. (2014) used the BIS to generate probabilities for burst events. However, there is no consensus on the issue of how to represent a leakage event, and currently, the rules that have been used to raise the alarm are determined intuitively by researchers.

Detection time (DT) or average detection time (ADT) describes the time duration between the start time of leakage and the time when a method successfully raised the alarm. It is important to raise the alarm as early as possible. In a real-life dataset, the DT could be difficult to be determined, and it could only be inferred from the customers' contacts or maintenance history. Based on the summary provided in Table 5, it could be observed that prediction-classification methods could receive the quickest response time for burst detection. It is hard to draw a conclusion about which method is the best method for leakage detection since each method has their own advantages and disadvantages. Also, the performance of a leakage detection method may depend on the time of occurrence and the magnitude and types of leakage. Moreover, different evaluation criteria are used in those studies, and thus it becomes more difficult to compare the performance.

Leakage Localisation

Currently, there is no consensus about the evaluation of leakage localisation methods. Most methods show their results using graphical representations of the probable leak nodes or areas and the true leak locations. The visualisation of the leak map could provide an intuitive view of the accuracy of leak localisation. However, quantification metrics are needed so that different methods can be compared with each other. Graphical distance to real leak and pipeline distance to real leak are the two quantification metrics that have been used in literature, and currently, the accuracy could only achieve 200 m (see Table 5). In addition, for classification-based method, the classification accuracy for leakage zone localisation have been used for method evaluation.

Leakage localisation is important to reduce disruption to customers and traffic by identifying the leak's location as close as possible. However, it is clear that accuracy and effectiveness have a great potential for improvement. Furthermore, it is important to evaluate the localisation methods and assess their capability in real-life datasets. The accuracy of leakage localisation is affected by several factors such as: (1) the size and types of leakage; (2) location of the leakage; (3) calibration of the model; (4) number and location of the sensors. Therefore, it is obvious that more comprehensive evaluation criteria for leakage localisation are needed. Considering the burst events, Qi et al. (2018) proposed a methodology to investigate the capacity of pressure-based burst detection using several quantitative metrics: (1) undetectable nodes, represent the effectiveness of a detection method and provide information of the need for additional sensors; (2) undetectable demands at those undetectable nodes to assess the potential capacity of a method; (3) detection dimension that indicates the correlation between nodes and the entire pressure sensor distribution; (4) spatial partition that investigate the influence of each sensor considering the distance of the sensor to the leakage; (5) detectable threshold that represents the minimum detectable burst flow. It is important to assess these quantitative metrics for a localisation method in the future research, especially for a system equipped with a large number of sensors.

CHALLENGES AND LIMITATIONS

Although the advantages and disadvantages of some methods have been discussed in step 3 and step 4, there are some other issues in leakage detection methods that still need to be addressed.

Different Types of Leakage

Leaks could happen in all WDNs. Depending on the size of leakage, they can also be categorised as: 1) Background leaks (small flow rate, invisible), 2) Unreported leaks (moderate flow rate, invisible), 3) Reported leaks/bursts (high flow rate, visible above ground). The burst events can be easily detected due to a large amount of water loss, but it will have a negative impact on customer satisfaction and may also cause contamination intrusion (Wu and Liu 2017). Compared with bursts, background and unreported leaks can accumulate into greater water loss due to a longer time to awareness. Pre-detection/detection of background and unreported leaks is challenging since the magnitude of leaks is small. Jung and Lansey (2015) seem reached an accurate detection result for small magnitude burst events, but the results were generated using synthetic data. Furthermore, the proposed method used nearly 2,000 days of normal data for statistics calculation, and this kind of information usually is unavailable in real life.

As shown in Table 1 and Table 2, most studies, especially data-driven methods, are focused on burst detection and did not evaluate the detection ability for gradual leakage events that developed from incipient leakages to burst events. Unlike bursts that can cause variation in a relatively short period (a few hours) and can be characterised by sudden flow increase and sudden pressure drop, gradual leakages will not generate noticeable deviation in the beginning stage and can be more challenging to detect. In addition, gradual leakage can cause more damage if they remain undetected. However, to the best of the authors' knowledge, there is no literature currently that have addressed this issue, and an early warning system for gradual leakage events could be a topic in future research. The dataset created by BattleDIM (Vrachminis et al. 2020) modelled different types of leakage (including burst, gradual leakage and background leakage) to evaluate the performance of competitors' methods. The automatic meter readings (AMRs) provide valuable information for accurate demand calibration, and most of those methods are developed based on the well-calibrated hydraulic model (Steffelbauer et al. 2022, Marzola et al. 2022). However, it should be noted that most water companies do not equip smart water meters in real-life. Moreover, due to the complexity of BattleDIM, Marzola et al. (2022) used engineering judgement and visual inspection instead of automated detection.

Data Uncertainties

In general, WDNs display the same daily water demand pattern, an increase in the early morning and late afternoon during weekdays and a slightly different pattern during weekends. Leakage acts as a demand in the network and affects pressure and flow values, but it does not follow the consumption patterns. The uncertainties within the data have posed great difficulties for leak detection. Most methods did not consider the demand variation caused by weather or the demand variation caused by population increases. The model needs to be retrained regularly to adapt to the changing situation, which may be time-consuming. Moreover, how to differentiate the variation caused by leakage and by weather, holiday behaviour, human activities etc., is still an issue that needs to be addressed.

How to handle the spurious outliers is also a question that needs to be considered during leakage detection. The current statistical information can be twisted by the incoming outliers and make the baseline that represents the normal behaviour inaccurate. In order to reduce the influence caused by outliers, Ye and Fenner (2014) used the expectation maximum algorithm to assign different weights to the data to reduce the influence of spurious points. Wang et al. (2020) designed a feedback loop to replace the detected outliers with a more appropriate value. Therefore, a feedback control system that can combine the information of online detection results and model updating is needed. It is especially important for real-time leak detection methods to involve the automatic adjustment of parameters during failed conditions or develop a method that is robust to outliers.

Hydraulic Model Uncertainties

Hydraulic model-based methods are influenced by uncertainties in both the model and the measurements. In the context of the hydraulic model, several simplifications in modelling will cause an unrealistic representation of the WDN, such as: 1) pipes that are considered not essential will be removed since it is computationally impractical to model all pipes of a large WDN; 2) water demands are aggregated at junctions during the modelling process, but in real life, water usage happened along pipes; 3) WDN input parameters contain uncertainties, such as pipe roughness, emitter coefficient. In addition, the accuracy of the hydraulic model could be influenced by uncertainty in nodal demands, measurements errors, etc. Thus, the residual between actual measurements and model output can be different from zero, even in the absence of leaks. It is important to consider these uncertainties and reduce the impact of these uncertainties so that a robust leak localisation method can be developed. Cugueró-Escofet et al. (2015) studied the effect of demand uncertainty on the ability of localisation methods, and Blesa and Pérez (2018) proposed a method of modelling the effect of these uncertainties on model-based localisation methods. The results suggested that future works should consider the uncertainty in the nominal value of the leak, inflow, and sensor measurements.

Unrealistic Assumptions

Most of the data-driven methods have a very important hidden assumption is that the historical data used for model training doesn't contain any leakage events and is under normal operation. This is because the accurate parameter estimation of the health situation of a WDN is needed to distinguish the abnormal event. Romano et al. (2014) used SPC for historical data selection to ensure the quality of training data. The accuracy of SPC is limited, and many normal data could be deleted even there are no leakage events. Another frequently used assumption is that the system operation stage is assumed stable during detection because it changes the behaviour of monitoring data. For example, if there is a pump station located downstream of a DMA that pumps water to a water tank, the pumping station will also cause sudden water outflow in the system, and the impact of the sudden pumping flow on the data pattern will be the same as a burst event. Future research should consider how to detect burst or leakage events even under the changing operation situation so that methods could be more robust in real-life applications.

Most hydraulic model-based methods use some unrealistic assumptions for real-world deployment such as: 1) assuming no measurement error, modelling error, calibration error, etc. 2) assuming that customer behaviour doesn't change; 3) modelling leakage events as aggregated demands at junctions or nodes, but in reality, the majority of leaks happen on pipes; 4) assuming that there are no uncertainties in pipe roughness coefficient, which may vary depending on the pipe materials, age, or encrusted materials on the pipe walls. Therefore, hydraulic model-based methods are targeted at finding leaks that occur after calibration. For leaks that have not been correctly identified and located, pipeline roughness values are often misadjusted to compensate for the head loss caused by those unidentified leaks (Wu et al. 2010a). Currently, model-based methods still have room for improvement and have not reached the maturity for mainstream adoption (Sophocleous et al. 2019). Thus, the question of how to develop a robust method to overcome these assumptions is not easily solved.

Spatial Relation of Sensors

Currently, very few studies (Wu et al. 2018a; b) have considered the data from multiple sensors at the same time, but developing a model for every sensor and the decision are made separately. Mounce et al. (2003) proposed a data fusion technology to fuse the information from different sensors and consider the cascading effect between each DMA. Multivariate methods which have been widely applied in fault detection of smart-grid (Zhou et al. 2019b) have a bright future in leakage detection. Compared with the univariate methods, multivariate analysis-based methods can consider the correlations and, therefore, should provide a more efficient detection performance (Ni et al. 2020).

From Table 1, it can be observed that the application of data-driven methods is mainly focused on leakage identification. However, it is worth exploring the application of data-driven methods for leakage localisation. Since the reference behaviour can be described by a prediction model or statistical characteristics, the influence of leakage to each sensor can be estimated by the deviance between leak data and reference value. The influence of leakage is relative to the distance between leakage and the sensor, and by quantifying the deviance value, the possible leakage area can be determined. Following this idea, Wu et al. (2018) identified the approximate location information by calculating the abnormality degree of each pressure sensor. Soldevila et al. (2019) have proposed data-driven leakage localisation methods by comparing the pressure map estimated by Kriging spatial interpolation, and Bayesian reasoning is applied to consider the temporal evolution to improve the accuracy of leakage localisation.

CONCLUSIONS

Leakage affects the majority of water utilities in developed and developing countries. Many efforts have been made to reduce leakage in water distribution systems. Leakage management has five aspects: prevention, assessment, control, detection, and localization and repair. In this article, the systematic review of leakage detection methods provides a compendium of information on existing technologies, their main implementation steps and issues of concern for each step, their suitability for different cases studies and their advantages and disadvantages.

A framework was developed to assess these methods. The framework has the following main steps: 1) definition of the main goal of the study, 2) data collection, 3) pre-processing of data, 4) data analysis and 5) method evaluation. Leakage detection methods based on pressure and flow data were categorized into hydraulic model-based and data-driven approaches. For leakage identification, data-driven methods can take advantage of a large amount of monitoring data and explore more valuable information. For leakage localisation, hydraulic model-based methods could take into account pressure data from multiple sensors and provide more accurate localisation. For each method, the advantages and disadvantages are provided in this paper.

In the future, researchers need to consider the limitations of current methodologies. Firstly, almost all papers focus on burst events, but it is crucial to develop an early warning system to detect gradual leakage before it causes obvious disruption and causes more water loss. Secondly, the uncertainties in both monitoring data and hydraulic models have impeded the application of leakage detection methods, and methods that can be robust to these uncertainties are needed in the future. Furthermore, it is important for leak detection methods to involve the automatic adjustment of parameters to achieve better real-time performance. Thirdly, it is important to realise the assumptions that have been made when developing the method, and real-life scenarios are the final goal that makes fewer assumptions. Fourthly, an information interaction system that can consider the information from multiple sensors is needed so that decision making could be made comprehensively instead of independently. Finally, a more comprehensive leakage detection evaluation method needs to be developed.

DATA AVAILABILITY STATEMENTS

No data, models, or code were generated or used during the study (e.g. opinion or dateless paper).

ACKNOWLEDGEMENT

The first author is funded by the China Scholarship Council (No. 202006370080), and the work is supported by a Royal Academy of Engineering Industrial Fellowship to resource Raziyeh Farmani's involvement (IF\192057).

REFERENCES

Abdulshaheed, A., Mustapha, F., and Ghavamian, A. (2017). "A Pressure-Based Method for Monitoring Leaks in a Pipe Distribution System: A Review." *Renewable and Sustainable Energy Reviews*, 69, 902–911.

Ahn, J., and Jung, D. (2019). "Hybrid Statistical Process Control Method for Water Distribution Pipe Burst Detection." *Journal of Water Resources Planning and Management*, 145(9), 06019008.

Aksela, K., Aksela, M., and Vahala, R. (2009). "Leakage Detection in a Real Distribution Network Using a SOM." *Urban Water Journal*, 6(4), 279–289.

Alla, S., and Adari, S. K. (2019). *Beginning Anomaly Detection Using Python-Based Deep Learning*. New Jersey: Apress.

Bakker, M., Vreeburg, J. H. G., Van De Roer, M., and Rietveld, L. C. (2014). "Heuristic Burst Detection Method Using Flow and Pressure Measurements." *Journal of Hydroinformatics*, 16(5), 1194–1209.

Bakker, M., Vreeburg, J. H. G., van Schagen, K. M., and Rietveld, L. C. (2013). "A Fully Adaptive Forecasting Model for Short-Term Drinking Water Demand." *Environmental Modelling and Software*, 48, 141–151.

Berglund, A., Areti, V. S., Brill, D., and Mahinthakumar, G. (Kumar). (2017). "Successive Linear Approximation Methods for Leak Detection in Water Distribution Systems." *Journal of Water Resources Planning and Management*, 143(8), 04017042.

Blesa, J., and Pérez, R. (2018). "Modelling Uncertainty for Leak Localization in Water Networks." *IFAC-PapersOnLine*, 51(24), 730–735.

Bohorquez, J., Alexander, B., Simpson, A. R., and Lambert, M. F. (2020). "Leak Detection and Topology Identification in Pipelines Using Fluid Transients and Artificial Neural Networks." *Journal of Water Resources Planning and Management*, 146(6), 1–11.

Chan, T. K., Chin, C. S., Member, S., and Zhong, X. (2018). "Review of Current Technologies and Proposed Intelligent Methodologies for Water Distributed Network Leakage Detection." *IEEE Access*, IEEE, 6, 78846–78867.

Colombo, A. F., Lee, P., and Karney, B. W. (2009). “A Selective Literature Review of Transient-Based Leak Detection Methods.” *Journal of Hydro-Environment Research*, 2(4), 212–227.

Cugueró-Escofet, P., Blesa, J., Pérez, R., Cugueró-Escofet, M., and Sanz, G. (2015). “Assessment of a Leak Localization Algorithm in Water Networks Under Demand Uncertainty.” *IFAC-PapersOnLine*, 28(21), 226–231.

Denyer, D., and Tranfield, D. (2006). “Using Qualitative Research Synthesis to Build an Actionable Knowledge base.” *Management decision*.

El-Zahab, S., Asaad, A., Mohammed Abdelkader, E., and Zayed, T. (2016). “Collective Thinking Approach for Improving Leak Detection Systems.” *Smart Water*, 2(1), 1–10.

Farah, E., and Shahrour, I. (2017). “Leakage Detection Using Smart Water System: Combination of Water Balance and Automated Minimum Night Flow.” *Water Resources Management*, 31(15), 4821–4833.

Farley, B., Mounce, S. R., and Boxall, J. B. (2013). “Development and Field Validation of a Burst Localization Methodology.” *Journal of Water Resources Planning and Management*, 139(6), 604–613.

Geiger, I. G. (2005). *Principles of Leak Detection*. KROHNE oil and gas 399.

Giorgio Bort, C. M., Righetti, M., and Bertola, P. (2014). “Methodology for Leakage Isolation Using Pressure Sensitivity and Correlation Analysis in Water Distribution Systems.” *Procedia Engineering*, 89, 1561–1568.

Gupta, A., and Kulat, K. D. (2018). “A Selective Literature Review on Leak Management Techniques for Water Distribution System.” *Water Resources Management*, 32(10), 3247–3269.

Han, Z., Zhao, J., Leung, H., Ma, K. F., and Wang, W. (2019). “A Review of Deep Learning Models for Time Series Prediction. ” *IEEE Sensors Journal*, 21(6), 7833–7848.

Harrou, F., Sun, Y., S. Hering, A., Madakyaru, M., and Dairi, A. (2020). *Statistical Process Monitoring using Advanced Data-Driven and Deep Learning Approaches*. Elsevier.

Hu, Z., Chen, B., Chen, W., Tan, D., and Shen, D. (2021a). “Review of model-based and data-driven approaches for leak detection and location in water distribution systems.” *Water Supply*, 21(7), 3282–3306.

Hu, X., Han, Y., Yu, B., Geng, Z., and Fan, J. (2021b). "Novel Leakage Detection and Water Loss Management of Urban Water Supply Network Using Multiscale Neural Networks." *Journal of Cleaner Production*, 278, 123611.

Huang, P., Zhu, N., Hou, D., Chen, J., Xiao, Y., Yu, J., Zhang, G., and Zhang, H. (2018). "Real-Time Burst Detection in District Metering Areas in Water Distribution System Based on Patterns of Water Demand with Supervised Learning." *Water (Switzerland)*, 10(12), 1–16.

Ismail, M. I. M., Dziyauddin, R. A., Salleh, N. A. A., Muhammad-Sukki, F., Bani, N. A., Izhar, M. A. M., and Latiff, L. A. (2019). "A Review of Vibration Detection Methods Using Accelerometer Sensors for Water Pipeline Leakage." *IEEE Access*, 7, 51965--51981.

Javadiha, M., Blesa, J., Soldevila, A., and Puig, V. (2019). "Leak Localization in Water Distribution Networks Using Deep Learning." *2019 6th International Conference on Control, Decision and Information Technologies, CoDIT 2019*, 1426–1431.

Jung, D., Kang, D., Liu, J., and Lansey, K. (2015). "Improving the Rapidity of Responses to Pipe Burst in Water Distribution Systems: A Comparison of Statistical Process Control Methods." *Journal of Hydroinformatics*, 17(2), 307–328.

Jung, D., and Lansey, K. (2015). "Water Distribution System Burst Detection Using a Nonlinear Kalman Filter." *Journal of Water Resources Planning and Management*, 141(5), 04014070.

Kang, D., and Lansey, K. (2014). "Novel Approach to Detecting Pipe Bursts in Water Distribution Networks." *Journal of Water Resources Planning and Management*, 140(1), 121–127.

Keramat, A., Wang, X., Louati, M., Meniconi, S., Brunone, B., and Ghidaoui, M. S. (2019). "Objective Functions for Transient-Based Pipeline Leakage Detection in a Noisy Environment: Least Square and Matched-Filter." *Journal of Water Resources Planning and Management*, 145(10), 04019042.

Loureiro, D., Amado, C., Martins, A., Vitorino, D., Mamade, A., and Coelho, S. T. (2016). "Water Distribution Systems Flow Monitoring and Anomalous Event Detection: A Practical Approach." *Urban Water Journal*, Taylor & Francis, 13(3), 242–252.

Marzola, I., Mazzoni, F., Alvisi, S., and Franchini, M. (2022). "Leakage Detection and Localization in a Water Distribution Network through Comparison of Observed and Simulated Pressure Data." *Journal of Water Resources Planning and Management*, 148(1), 1–11.

Misiunas, D., Vitkovský, J., Olsson, G., Lambert, M., and Simpson, A. (2006). "Failure Monitoring in Water Distribution Networks." *Water Science and Technology*, 53(4–5), 503–511.

Mounce, S. R., Boxall, J. B., and Machell, J. (2007). "An Artificial Neural Network/Fuzzy Logic System for DMA Flow Meter Data Analysis Providing Burst Identification and Size Estimation." *Proceedings of the Combined International Conference of Computing and Control for the Water Industry, CCWI2007 and Sustainable Urban Water Management, SUWM2007*, (January), 313–320.

Mounce, S. R., Boxall, J. B., and Machell, J. (2010). "Development and Verification of an Online Artificial Intelligence System for Detection of Bursts and Other Abnormal Flows." *Journal of Water Resources Planning and Management*, 136(3), 309–318.

Mounce, S. R., Day, A. J., Wood, A. S., Khan, A., Widdop, P. D., and Machell, J. (2002). "A Neural Network Approach to Burst Detection." *Water Science and Technology*, 45(4–5), 237–246.

Mounce, S. R., Khan, A., Wood, A. S., Day, A. J., Widdop, P. D., and Machell, J. (2003). "Sensor-Fusion of Hydraulic Data for Burst Detection and Location in a Treated Water Distribution System." *Information Fusion*, 4(3), 217–229.

Mounce, S. R., and Machell, J. (2006). "Burst Detection Using Hydraulic Data from Water Distribution Systems with Artificial Neural Networks." *Urban Water Journal*, 3(1), 21–31.

Mounce, S. R., Mounce, R. B., and Boxall, J. B. (2011). "Novelty Detection for Time Series Data Analysis in Water Distribution Systems Using Support Vector Machines." *Journal of Hydroinformatics*, 13(4), 672–686.

Mounce, S. R., Mounce, R. B., and Boxall, J. B (2012). "Identifying Sampling Interval for Event Detection in Water Distribution Networks." *Journal of Water Resources Planning and Management*, 138(2), 187–191.

Ni, F. T., Zhang, J., and Noori, M. N. (2020). "Deep Learning for Data Anomaly Detection and Data Compression of a Long-Span Suspension Bridge." *Computer-Aided Civil and Infrastructure Engineering*, 35(7), 685–700.

Okeya, I., Hutton, C., and Kapelan, Z. (2015). "Locating Pipe Bursts in a District Metered Area via Online Hydraulic Modelling." *Procedia Engineering*, 119(1), 101–110.

Palau, C. V., Arregui, F. J., and Carlos, M. (2012). "Burst Detection in Water Networks Using Principal Component Analysis." *Journal of Water Resources Planning and Management*, 138(1), 47–54.

Pérez, R., Puig, V., Pascual, J., Quevedo, J., Landeros, E., and Peralta, A. (2011). "Methodology for Leakage Isolation Using Pressure Sensitivity Analysis in Water Distribution Networks." *Control Engineering Practice*, 19(10), 1157–1167.

Perez, R., Sanz, G., Puig, V., Quevedo, J., Cuguero Escofet, M. A., Nejjari, F., Meseguer, J., Cembrano, G., Mirats Tur, J. M., and Sarrate, R. (2014). "Leak Localization in Water Networks: A Model-Based Methodology Using Pressure Sensors Applied to a Real Network in Barcelona." *IEEE Control Systems*, 34(4), 24–36.

Porwal, S., Akbar, S. A., and Jain, S. C. (2017). "Leakage Detection and Prediction of Location in a Smart Water Grid using SVM Classification." *International Conference on Energy, Communication, Data Analytics and Soft Computing*, IEEE, 3288–3292.

Puust, R., Kapelan, Z., Savic, D. A., and Koppel, T. (2010). "A Review of Methods for Leakage Management in Pipe Networks." *Urban Water Journal*, 7(1), 25–45.

Qahtani, T. Al, Yaakob, M. S., Yidris, N., and Sulaiman, S. (2020). "A Review on Water Leakage Detection Method in the Water Distribution Network." *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, 68(2), 152–163.

Qi, Z., Zheng, F., Guo, D., Maier, H. R., Zhang, T., Yu, T., and Shao, Y. (2018). "Better Understanding of the Capacity of Pressure Sensor Systems to Detect Pipe Burst within Water Distribution Networks." *Journal of Water Resources Planning and Management*, 144(7), 04018035.

Romano, M., Kapelan, Z., and Savić, D. A. (2011). "Real-Time Leak Detection in Water Distribution Systems." *Water Distribution Systems Analysis 2010*, American Society of Civil Engineers, Reston, VA, 1074–1082.

Romano, M., Kapelan, Z., and Savić, D. A. (2014). "Automated Detection of Pipe Bursts and Other Events in Water Distribution Systems." *Journal of Water Resources Planning and Management*, 140(4), 457–467.

Romero, L., Blesa, J., Puig, V., and Cembrano, G. (2022). "Clustering-Learning Approach to the Localization of Leaks in Water Distribution Networks." *Journal of Water Resources Planning and*

Management, 148(5), 1–11.

Sanz, G., Pérez, R., Kapelan, Z., and Savic, D. (2016). “Leak Detection and Localization through Demand Components Calibration.” *Journal of Water Resources Planning and Management*, 142(2), 04015057.

Soldevila, A., Blesa, J., Fernandez-Canti, R. M., Tornil-Sin, S., and Puig, V. (2019). “Data-Driven Approach for Leak Localization in Water Distribution Networks Using Pressure Sensors and Spatial interpolation.” *Water (Switzerland)*, 11(7).

Soldevila, A., Blesa, J., Tornil-Sin, S., Duviella, E., Fernandez-Canti, R. M., and Puig, V. (2016). “Leak Localization in Water Distribution Networks Using a Mixed Model-Based/Data-Driven Approach.” *Control Engineering Practice*, 55, 162–173.

Sophocleous, S. (2019). “Development of the Next Generation of Water Distribution Network Modelling Tools Using Inverse Methods.”

Sophocleous, S., Savić, D., and Kapelan, Z. (2019). “Leak Localization in a Real Water Distribution Network Based on Search-Space Reduction.” *Journal of Water Resources Planning and Management*, 145(7), 04019024.

Sophocleous, S., Savic, D., Kapelan, Z., Gilbert, C., and Sage, P. (2018). “Leak Detection and Localization Based on Search Space Reduction and Hydraulic Modelling.” *WDSA / CCWI Joint Conference Proceedings*.

Steffelbauer, D. B., Deuerlein, J., Gilbert, D., Abraham, E., and Piller, O. (2022). “Pressure-Leak Duality for Leak Detection and Localization in Water Distribution Systems.” *Journal of Water Resources Planning and Management*, 148(3), 1–13.

Vrachimis, S. G., Eliades, D. G., Taormina, R., Ostfeld, A., Kapelan, Z., Liu, S., Kyriakou, M., Pavlou, P., Qiu, M., and Polycarpou, M. M. (2020). “BattLeDIM : Battle of the Leakage Detection and Isolation Methods. ” *2nd International CCWI/WDSA Joint Conference*.

Wagner, J.M., Shamir, U., and Marks, D.H. (1988). “Water Distribution Reliability: Simulation Methods. ” *Journal of Water Resources Planning and Management*, 114(3), 276–294.

Wang, X., Guo, G., Liu, S., Wu, Y., Xu, X., and Smith, K. (2020). “Burst Detection in District Metering Areas Using Deep Learning Method.” *Journal of Water Resources Planning and Management*, 146(6), 1–12.

Wu, Y., and Liu, S. (2017). “A Review of Data-Driven Approaches for Burst Detection in Water Distribution Systems.” *Urban Water Journal*, Taylor & Francis, 14(9), 972–983.

Wu, Y., Liu, S., and Asce, A. M. (2020). “Burst Detection by Analyzing Shape Similarity of Time Series Subsequences in District Metering Areas.” *Journal of Water Resources Planning and Management*, 146(1), 1–12.

Wu, Y., Liu, S., Smith, K., and Wang, X. (2018a). “Using Correlation between Data from Multiple Monitoring Sensors to Detect Bursts in Water Distribution Systems.” *Journal of Water Resources Planning and Management*, 144(2), 04017084.

Wu, Y., Liu, S., and Wang, X. (2018b). “Distance-Based Burst Detection Using Multiple Pressure Sensors in District Metering Areas.” *Journal of Water Resources Planning and Management*, 144(11), 06018009.

Wu, Y., Liu, S., Wu, X., Liu, Y., and Guan, Y. (2016). “Burst Detection in District Metering Areas Using a Data Driven Clustering Algorithm.” *Water Research*, 100, 28–37.

Wu, Z. Y., Sage, P., and Turtle, D. (2010a). “Pressure-Dependent Leak Detection Model and Its Application to a District Water System.” *Journal of Water Resources Planning and Management*, 136(1), 116–128.

Wu, Z. Y., Burrows, R., Moorcroft, J., Croxton, N., and Limanond, S. (2010b). “Pressure-Dependent Leakage Detection Method Compared with Conventional Techniques.” *Water Distribution Systems Analysis 2010*, 1083–1092.

Wu, J., Ma, D., and Wang, W. (2022). “Leakage Identification in Water Distribution Networks Based on XGBoost Algorithm.” *Journal of Water Resources Planning and Management*, 148(3), 1–13.

Xu, Q., Liu, R., Chen, Q., and Li, R. (2014). “Review on Water Leakage Control in Distribution Networks and the Associated Environmental Benefits.” *Journal of Environmental Sciences*, The Research Centre for Eco-Environmental Sciences, Chinese Academy of Sciences, 26(5), 955–961.

Xu, Z., Ying, Z., Li, Y., He, B., and Chen, Y. (2020). “Pressure Prediction and Abnormal Working Conditions Detection of Water Supply Network Based on LSTM.” *Water Science and Technology: Water Supply*, 20(3), 963–974.

Ye, G., and Fenner, R. A. (2011). “Kalman Filtering of Hydraulic Measurements for Burst Detection in Water Distribution Systems.” *Journal of Pipeline Systems Engineering and Practice*, 2(1), 14–22.

Ye, G., and Fenner, R. A. (2014). “Weighted Least Squares with Expectation-Maximization Algorithm for Burst Detection in U.K. Water Distribution Systems.” *Journal of Water Resources Planning and Management*, 140(4), 417–424.

Zaman, D., Tiwari, M. K., Gupta, A. K., and Sen, D. (2020). “A Review of Leakage Detection Strategies for Pressurised Pipeline in Steady-State.” *Engineering Failure Analysis*, 109, 104264.

Zhang, Q., Wu, Z. Y., Zhao, M., Qi, J., Huang, Y., and Zhao, H. (2016). “Leakage Zone Identification in Large-Scale Water Distribution Systems Using Multiclass Support Vector Machines.” *Journal of Water Resources Planning and Management*, 142(11), 04016042.

Zhou, X., Tang, Z., Xu, W., Meng, F., Chu, X., Xin, K., and Fu, G. (2019a). “Deep Learning Identifies Accurate Burst Locations in Water Distribution Networks.” *Water Research*, 166, 115058.

Zhou, Y., Arghandeh, R., Zou, H., and Spanos, C. J. (2019b). “Nonparametric Event Detection in Multiple Time Series for Power Distribution Networks.” *IEEE Transactions on Industrial Electronics*, 66(2), 1619–1628.

Table 1. Summary of steps 0-1 of data-driven leakage detection studies

Category	Reference	Goal (Step 0)	Data Collection (Step 1)				
			Case Study	Data Type	Frequency	Leak Type	Leak Size
SPC-based methods	Misiunas et al. 2006	Identification/ Localisation	1 simulated dataset	Flow	1 min	Burst	5-10 L/s
	Palau et al. 2012	Identification	1 historical dataset	Flow	5 min	Burst	Around 5% of the average flow
	Jung et al. 2015	Identification	2 simulated datasets, 1 historical dataset	Flow, Pressure	5 min	Burst	0.13%-0.72% of the average flow
	Loureiro et al. 2016	Identification	1 historical dataset	Flow	15 min	Burst	--
	Ahn and Jung 2019	Identification	1 simulated dataset	Flow	5 min	Burst	0.1%-3.3% of the mean total demand
Prediction-classification methods	Mounce et al. 2002	Identification	1 engineered test dataset	Flow, Pressure	15 min	Burst	--
	Mounce et al. 2007	Identification	1 engineered test dataset	Flow	15 min	Burst	5-7 L/s
	Mounce et al. 2010	Identification	1 historical dataset	Flow	15 min	Burst/Leak	9%-32% of the average demand
	Romano et al. 2011, 2014	Identification	1 engineered test dataset	Flow	15 min	Burst	5%-16% of the average inflow
	Mounce et al. 2011	Identification	1 engineered test dataset, 1 historical dataset	Flow, Pressure	15 min	Burst	6%-12% of the average daily maximum flow
	Ye and Fenner 2011, 2014	Identification	1 engineered test dataset, 1 historical dataset	Flow, Pressure	15 min	Burst/Leak	1-5 L/s
	Bakker et al. 2014	Identification	1 historical dataset	Flow, Pressure	5 min	Burst	150 m^3/h for largest area, 7 m^3/h for smallest area
	Jung and Lansey 2015	Identification	1 simulated dataset	Flow, Pressure	5 min	Burst	0.3%–7.0% of the mean total demand
	Wang et al. 2020	Identification	1 engineered test dataset	Flow	5 min	Burst	2.8%-14% of the average inflow
	Xu et al. 2020	Identification	1 engineered test dataset	Flow, Pressure	5 min	Burst	--
Clustering-based methods	Aksela et al. 2009	Identification	1 historical dataset	Flow	1 hour	Leak	--
	Wu et al. 2016	Identification	1 engineered test dataset	Flow	5 min	Burst	13.3%-23.1% of current inflow
	Wu et al. 2018a	Identification	1 engineered test dataset	Pressure	5 min	Burst	13.3%-23.1% of current inflow
	Huang et al. 2018	Identification	1 engineered test dataset	Flow	15 min	Burst	10%-20% of the average inflow
	Wu et al. 2020b	Identification	1 simulated dataset, 1 engineered test dataset	Flow	5 min, 15 min	Burst	6%–12% of the average inflow

-- represents that the information did not provide in the paper

Table 2. Summary of steps 0-1 of hydraulic model-based leakage detection studies

Category	Reference	Goal (Step 0)	Data Collection (Step 1)				
			Case Study	Data Type	Frequency	Leak Type	Leak Size
Calibration-based methods	Misiunas et al. 2006	Identification, Localisation	1 simulated dataset	Pressure	1 min	Burst	5-10 L/s
	Wu et al. 2010b	Localisation	1 engineered test dataset, 1 historical dataset	Pressure	30 min	Leak	--
	Sanz et al. 2016	Identification, Localisation	1 simulated dataset	Flow, Pressure	10 min	Burst	2.5%-13% of total consumption
	Berglund et al. 2017	Localisation	3 simulated datasets	Pressure	1 hour	Leak	Less than 0.5% of the total inflow
	Sophocleous et al. 2018, 2019	Localisation	1 simulated dataset, 1 historical dataset	Flow, Pressure	15 min	Burst	5%-50% of the inlet flow
Sensitivity analysis-based methods	Farley et al. 2013	Localisation	1 engineered test dataset	Flow, Pressure	15 min	Burst	--
	Kang and Lansey 2014	Identification, Localisation	1 simulated dataset	Flow, Pressure	1 hour	Burst	Emitter coefficient of 0.1
	Perez et al. 2014	Localisation	1 engineered test dataset	Flow, Pressure	10 min	Leak	About 5.6 L/s
	Okeya et al. 2015	Localisation	1 simulated dataset	Flow, Pressure	15 min	Burst	5%-50% of the average demand
	Giorgio Bort et al. 2014	Localisation	1 simulated dataset	Pressure	--	Leak	--
	Steffelbauer et al. 2022	Identification, Localisation	1 simulated dataset	Flow, Pressure, AMR	5 min	Burst and leak	5-30 m^3/h
Classification-based methods	Soldevila et al. 2016	Localisation	3 simulated datasets	Flow, Pressure	1 hour	Leak	0.84%-2.51% of the total demand
	Zhang et al. 2016	Localisation	2 simulated datasets	Pressure	--	Leak	Around 3% of the average demand
	Porwal et al. 2017	Localisation	1 simulated dataset	Flow, Pressure	30 min	Leak	Emitter coefficient of 0.005-0.1
	Zhou et al. 2019a	Localisation	2 simulated datasets	Pressure	15 min	Burst	Intensity coefficient of 10%-30%
	Hu et al. 2021b	Localisation	1 engineered test dataset	Flow, Pressure	--	Leak	10-38 L/s
	Romero et al. 2022	Localisation	1 historical dataset	Pressure	2 min	Leak	1.15 L/s
	Wu et al. 2022	Localisation	2 simulated datasets	Pressure	--	Leak	0-25 L/s

-- represents that the information did not provide in the paper

Table 3. Summary of steps 2-4 of data-driven leakage detection studies

Reference	Data Preparation (Step 2)	Data Analytics (Step 3)		Evaluation metrics (Step 4)
SPC-Based Methods				
Misiunas et al. 2006	RLS (Denoising)	CUSUM		DT
Palau et al. 2012	Mean centering and scaling, PCA	Hotelling T^2, DMOD		Detection effectiveness
Jung et al. 2015	Normalisation	WEC, CUSUM, EWMA, M-CUSUM, M-EWMA, Hotelling T2		ADT, TPR, NF
Loureiro et al. 2016	Moving average, data correction, normalisation	Modified Shewhart chart		TPR, FPR
Ahn and Jung 2019	Normalisation	Hybrid method of WEC and CUSUM		ADT, TPR, FPR
Prediction-Classification Methods				
Reference	Data Preparation (Step 2)	Data Analytics (Step 3)		Evaluation metrics (Step 4)
		Prediction model	Classification	
Mounce et al. 2002	Data correction, normalisation, reformatting	MDN	A classification module	--
Mounce et al. 2007, 2010	Data correction, normalisation, reformatting	ANN	FIS	--
Romano et al. 2011, 2014	SPC (Data selection), data correction, WT (de-noising)	ANN	SPC, BIS	DT, AUC
Mounce et al. 2011	Reformatting	SVR	Binomial event discriminator	--
Ye and Fenner 2011	Reformatting	KF	A user-defined threshold	--
Ye and Fenner 2014	Reformatting	Weighted least squares	A user-defined threshold	--
Bakker et al. 2014	--	Adaptive forecasting model (Bakker et al. 2013)	A user-defined threshold	DT, TPR, FPR, AUC
Jung and Lansey 2015	--	KF, NKF	CUSUM, Hotelling T^2	ADT, TPR, FPR
Wang et al. 2020	Reformatting	LSTM	Multithreshold classification based on time-varying z-score	DT, TPR, FPR
Xu et al. 2020	Linear interpolation (data correction), WT (data de-noising), normalisation	A parallel LSTM tandem deep neural network	A user defined value	DT, NF
Clustering-Based Methods				
Reference	Data Preparation (step 2)	Data Analysis (Step 3)		Evaluation metrics (Step 4)
		Similarity measure	Clustering	
Aksela et al. 2009	--	A leak function	SOM	
Wu et al. 2016	Reformatting	Euclidean distance	Clustering algorithm	TPR, FPR
Wu et al. 2018a	Data selection, reformatting, normalisation	Cosine distance	Clustering algorithm	TPR, FPR
Huang et al. 2018	Data selection	DTW	Random forest	TPR, FPR
Wu et al. 2020b	Normalisation, data selection, reformatting	Increase-state distance	ASS algorithm	TPR, NF, FPR

-- represents that the information did not provide in the paper

Table 4. Summary of steps 2-4 of hydraulic model-based leakage detection studies

Calibration-Based Methods				
Reference	Data Preparation (Step 2)	Data Analytics (Step 3)		Evaluation metrics (Step 4)
		Objective Function	Optimisation Algorithm	
Misiunas et al. 2006	De-noising (RLS), model calibration	Minimising the sum of difference squares	Trial-and-error	Comparison of candidate node and predict node
Wu et al. 2010b	Model calibration	1. Minimise the sum of difference squares 2. Minimise the sum of absolute differences 3. Minimise the maximum absolute difference	Genetic Algorithm (GA)	Geographic distance to leak
Sanz et al. 2016	Model calibration	Minimising the error in pressure and flow measurements	Least squares	Graphical distance to leak, Pipe distance to leak
Berglund et al. 2017	Model calibration	Minimising the sum of absolute pressure difference	LP, MILP	Comparison of candidate node and predict node
Sophocleous et al. 2018, 2019	Model calibration	1. single-leak: minimise the weighted sum of squared flow 2. n leak: minimise the weighted sum of squared differences for both pressure and flow	Search space reduction, GA	Geographic distance to leak

Sensitivity Analysis-Based Methods				
Reference	Data Preparation (Step 2)	Data Analytics (Step 3)		Evaluation metrics (Step 4)
		Sensitivity Analysis	Decision Making	
Farley et al. 2013	Model calibration	Jacobian sensitivity matrix	GA	--
Giorgio Bort et al. 2014	Model calibration	Sensitivity matrix	PCA, Least squares	--
Kang and Lansey 2014	Model calibration	Binarised sensitivity matrix	Statistical analysis	--
Perez et al. 2014	Reformatting, Model calibration	Sensitivity matrix	Biggest correlation values	Visualization of leak map
Okeya et al. 2015	Model calibration	Binarised matrix	Trial-and-error	Visualization of leak map
Steffelbauer et al. 2022	Model calibration based on a so-called dual approach	Jacobian sensitivity matrix	Highest pearson correlation sum	Geographic distance to leak

Classification-Based Methods				
Reference	Data Preparation (Step 2)	Data Analytics (Step 3)		Evaluation metrics (Step 4)
		Training Data	Classifier	
Soldevila et al. 2016	Model calibration, data generation, node grouping	Pressure residuals	kNN	Graphical distance to leak
Zhang et al. 2016	Model calibration, zone division (k-means), data generation	Pressure residuals	M-SVM	Visualization of leak map, classification accuracy
Porwal et al. 2017	Model calibration, data generation	Leakage and non-leakage scenario	SVM	Classification accuracy
Zhou et al. 2019a	Model calibration, data generation	Leakage and non-leakage scenario	FL-DenseNet	Visualization of leak map
Hu et al. 2021b	Model calibration	Pressure and flow residuals	DBSCAN-MFCN	Comparison of candidate node and predict node

Romero et al. 2022	Data generation, image encoding (GAF), zone division (GAC)	Pressure data	DNN	Pipe distance to leak
Wu et al. 2022	Data generation, zone division (FCM)	Pressure residuals	XGBoost	Classification accuracy

-- represents that the information did not provide in the paper

Table 5. Summary of performance evaluation of each method

Performance evaluation criteria	Results	References	Categories
		Leak identification	
Detection time (DT) or average detection time (ADT)	Within 5 minutes	Misiunas et al. 2006	SPC-based
	Around 2-12 hours	Jung et al. 2015	SPC-based
	1 hour 45 minutes average		Prediction-classification
	In most cases, within 15 minutes	Romano et al. 2011, 2014	Prediction-classification
	Less than 30 minutes	Bakker et al. 2014	Prediction-classification
	Around 1 – 5 hours	Jung and Lansey 2015	Prediction-classification
	Within 10 minutes	Wang et al. 2020	Prediction-classification
	Within 10 minutes	Xu et al. 2020	Prediction-classification
	Around 3-10 hours	Sanz et al. 2016	Calibration-based
True positive rate (TPR) Or detection probability (DP)	Around 55% - 78%	Jung et al. 2015	SPC-based
	Around 80% - 93%	Loureiro et al. 2016	SPC-based
	In the beset model 81%	Ahn and Jung 2019	SPC-based
	In the best case, 90%	Bakker et al. 2014	Prediction-classification
	In the best model, 98%	Jung and Lansey 2015	Prediction-classification
	In the best model, 100%	Wang et al. 2020	Prediction-classification
	In the best model, 71.43%	Wu et al. 2016	Clustering-based
	In the best model, 100%	Huang et al. 2018	Clustering-based
	In the best model, 90%	Wu et al. 2020b	Clustering-based
False Positive rate (FPR)	Around 10% - 16%	Loureiro et al. 2016	SPC-based
	Around 0 - 1% per day	Ahn and Jung 2019	SPC-based
	In the best cases, 2.1%	Bakker et al. 2014	Prediction-classification
	Around 0 – 1% per day	Jung and Lansey 2015	Prediction-classification
	In the best model, 0.41%	Wang et al. 2020	Prediction-classification
	Around 0.4% – 0.8%	Wu et al. 2016	Clustering-based
	In the best model, 0%	Huang et al. 2018	Clustering-based
	Around 5% - 10%	Wu et al. 2020b	Clustering-based
Number of false positives (NF)	1-18 false positives per day	Jung and Lansey 2015	Prediction-classification
	2 false positives per day	Xu et al. 2020	Prediction-classification
	50-100 false positives	Wu et al. 2020b	Clustering-based
AUC	0.88	Romano et al. 2011, 2014	Prediction-classification
	0.972 for larger bursts, 0.535 for all bursts	Bakker et al. 2014	Prediction-classification
		Leak localisation	
Geographic distance to leak	In most cases, around 200 m	Sanz et al. 2016	Calibration-based
	Within an area of 100 m radius	Wu et al. 2010b	Calibration-based
	In most cases, within 200 m	Sophocleous et al. 2018, 2019	Calibration-based
	Around 200 m	Soldevila et al. 2016	Classification-based
	Within 500 m	Porwal et al. 2017	Classification-based
Pipe distance to leak	In most cases, around 200-400 m	Sanz et al. 2016	Calibration-based
	In most cases, within 250 m	Steffelbauer et al. 2022	Sensitivity analysis-based
	Leakage scope is narrowed to the pipelines near the predicted leakage nodes	Zhang et al. 2016	Classification-based
	442 m	Romero et al. 2022	Classification-based
Classification accuracy	Around 80.78% - 99.25%	Zhang et al. 2016	Classification-based
	Around 40% - 90%	Porwal et al. 2017	Classification-based

	Around 85 % - 100%	Zhou et al. 2019	Classification-based
	Around 67.2% - 90.4%	Wu et al. 2022	Classification-based

Fig. 1. Classification of leakage management and leakage detection methods

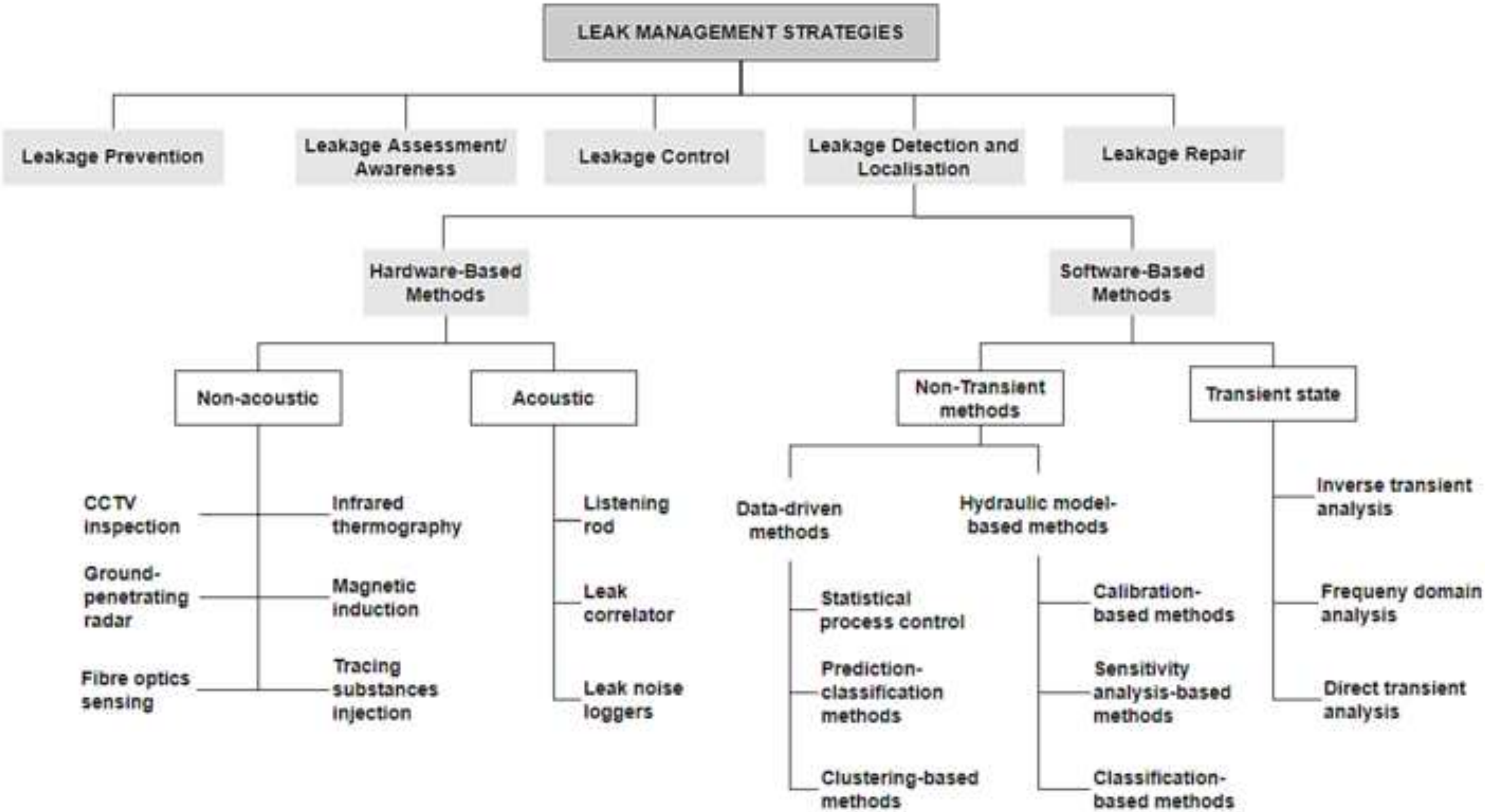


Fig. 2. The framework of data analytics for leakage detection

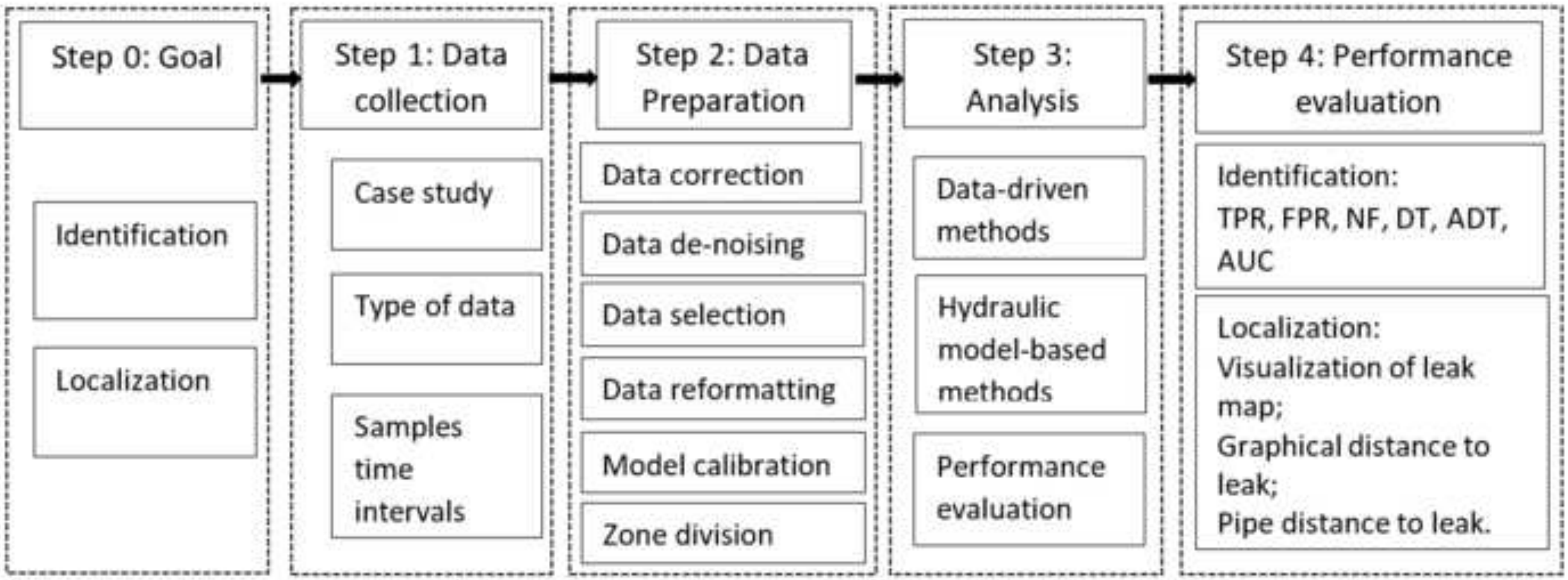


Fig. 3. Data analytic steps for data-driven leakage detection methods

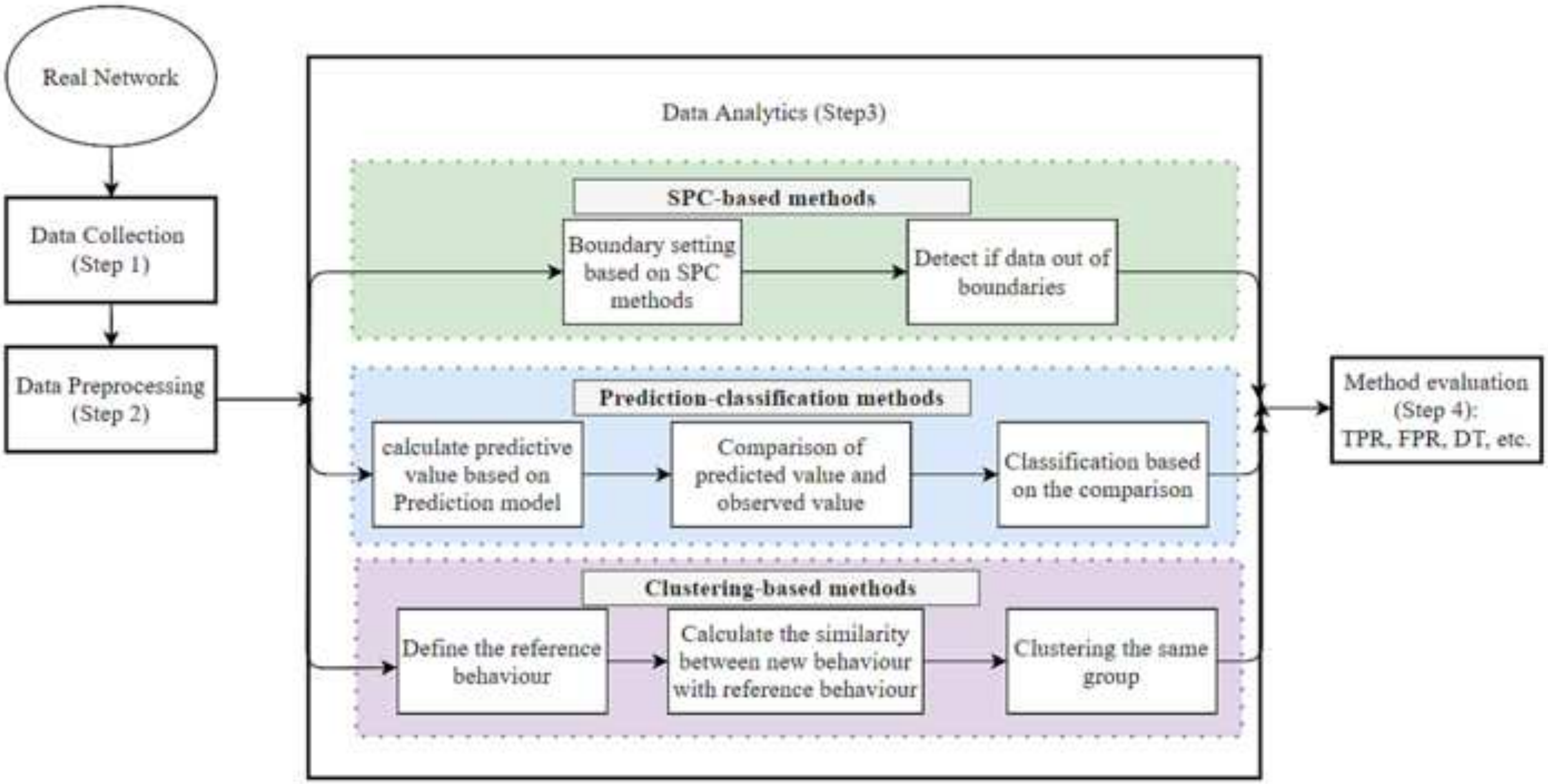
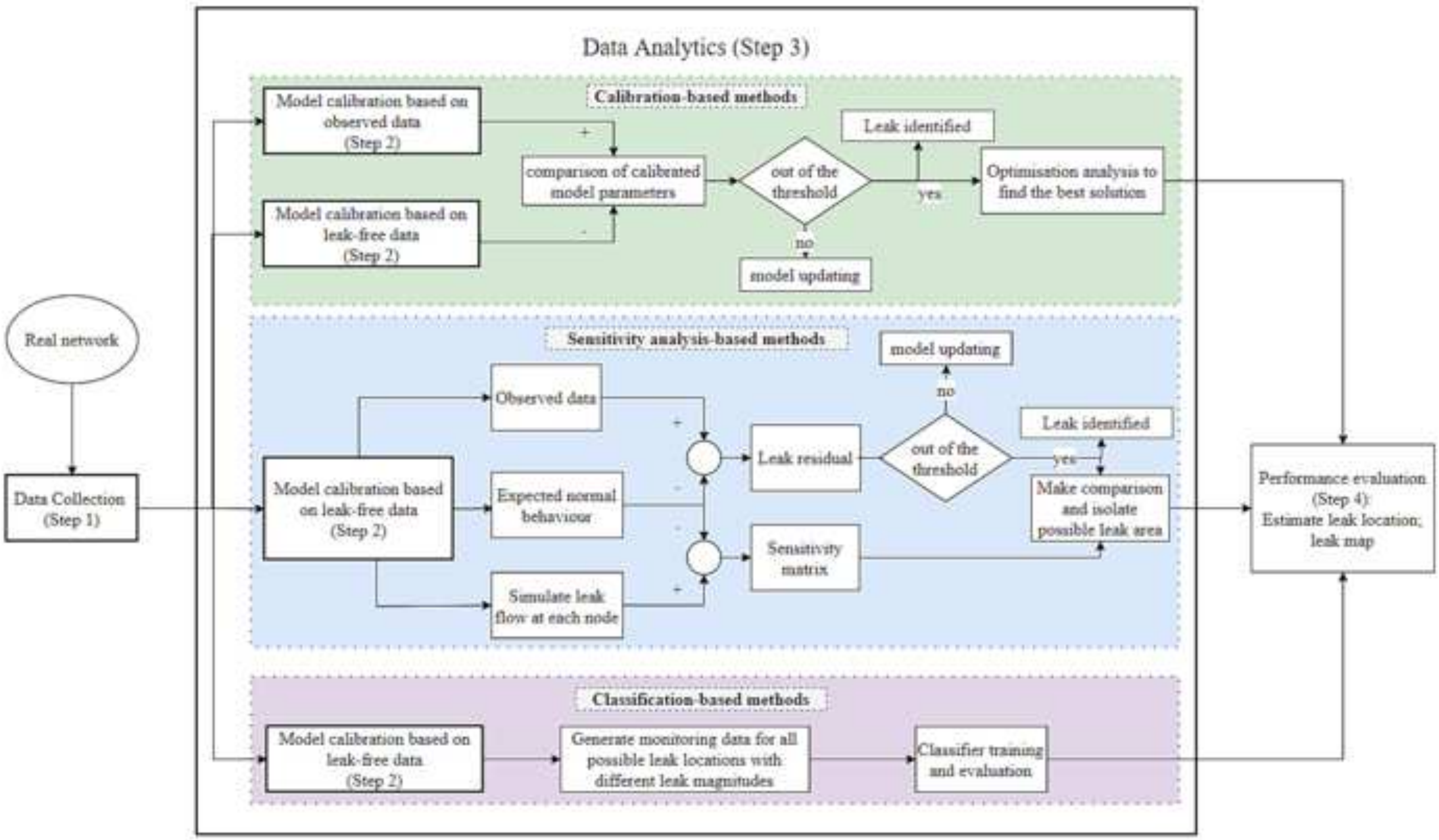


Fig. 4. Data analytic steps for hydraulic model-based leakage detection methods



List of Figures

Fig. 1. Classification of leakage management and leakage detection methods

Fig. 2. The framework of data analytics for leakage detection

Fig. 3. Data analytic steps for data-driven leakage detection methods

Fig. 4. Data analytic steps for hydraulic model-based leakage detection methods