

Article

Not peer-reviewed version

Optimal Pressure Sensor Deployment for Leak Identification in Water Distribution Networks

Guang Yang and [Hai Wang](#) *

Posted Date: 16 May 2023

doi: 10.20944/preprints202305.1158.v1

Keywords: leak identification; pressure sensor deployment; water distribution networks



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Optimal Pressure Sensor Deployment for Leak Identification in Water Distribution Networks

Guang Yang and Hai Wang *

School of Mechanical Engineering, Tongji University, No. 1239, Siping Road, Shanghai 200092, China; 862857086@qq.com

* Correspondence: wanghai@tongji.edu.cn;

Abstract: Pipe leakage is an inevitable phenomenon in water distribution networks (WDNs), leading to energy waste and economic damage. Leakage events can be reflected quickly by pressure values, and the deployment of pressure sensors is significant for minimizing the leakage ratio of WDNs. Concerning the restriction of realistic factors, including project budgets, available sensor installation locations, and sensor fault uncertainties, a practical methodology is proposed in this paper to optimize pressure sensor deployment for leak identification in terms of these realistic issues. Two indexes are identified to evaluate the leak identification ability, that is, detection coverage rate (*DCR*) and total detection sensitivity (*TDS*), and the principle is to determine priority to ensure an optimal *DCR* and retain the largest *TDS* with an identical *DCR*. Leakage events are generated by a model simulation and the essential sensors for maintaining the *DCR* are obtained by subtraction. In the event of a surplus budget, and if we suppose the partial sensors have failed, then we can determine the supplementary sensors that can best complement the lost leak identification ability. Moreover, a typical WDN Net3 is employed to show the specific progress, and the result shows that the methodology is largely appropriate for real projects.

Keywords: leak identification; pressure sensors deployment; water distribution networks

1. Introduction

A water supply network is an important element of civil infrastructure for meeting the needs of people's production and living water. Due to aging, corrosion, and human impacts on pipelines, pipeline leakage accidents occur frequently and result in a series of economic and social problems, such as water pollution and land subsidence [1]. According to the *Statistical Yearbook of Urban Water Distribution Networks* in China, the average leakage rate has decreased from 14.57% in 2017 to 12.68% in 2021 [2]. However, there is still a gap compared to developed European countries. The average leakage rates of most countries remain below 10% and even below 5% in Germany and the Netherlands [3]. China's Ministry of Housing and Urban–Rural Development and the NDRC (National Development and Reform Commission) jointly issued a notice for strengthening the leakage control of public water supply pipe networks, which requires that the leakage rates of urban public water supply pipe networks nationwide should be controlled to within 9% by 2025. With the developments in science and society, fine management requirements for the water supply industry have been proposed, and the leakage rates are expected to be further decreased to bring more social and economic benefits, with the exception of the DMA (District Metering Area) management, which deducts the metering losses [4].

Nowadays, the methods for leak identification can be roughly divided into three categories: equipment-based methods, transient-based methods, and data-driven methods [5,6]. The first methods relay information using hardware devices installed in the pipelines, and when leakage occurs, the devices can detect and determine the specific locations of the leaks. The representative equipment-based methods include sound-wave detection [7], the smart ball method [8], fiber optic cables [9], thermal imaging technology [10], etc. With IoT (Internet-of-Things) technologies[11], the transient-based methods and data-driven methods have been popular methods for leak location identification in practical engineering. The transient-based methods [12] and data-driven methods

[13] rely on real-time monitoring hydraulic data, with the latter being more seriously dependent on these data. The difference is that the transient-based methods have a fundamental physical model while the data-driven methods mainly utilize various machine learning algorithms, such as the genetic algorithm [14], support vector machine [15], and neural networks [16,17], to predict pressures/flows compared with observed values [18]. Hydraulic data in water supply network objectively reflect the operating status of an entire network and provide the basis for various leak identification methods. Pressure sensors have the advantages of being low cost and non-invasive, with rapid response capabilities [19]. However, the overuse of pressure sensors will increase the cost burden and consume redundant manpower and material resources. Therefore, optimizing the pressure sensor arrangement in a water supply network is essential for improving operations management on the promise of a limited budget.

Compared to the studies on leakage detection and location, there are few researchers who have dealt with pressure sensor layout optimization focused on leak identification. Generally, the limited studies on optimal pressure sensor deployment in WDNs can be classified into those that used the sensitivity analysis method [20,21], the cluster analysis method [22], the correlation analysis method [23], and the optimization algorithm [24]. In terms of the optimization algorithm, it has the advantage of systematicity and reliability. However, the position of sensors in a large pipeline network is an explosive combination problem, and a more efficient and reasonable optimization method is necessitated to address this challenge. Juan Li et al. (2019) [25] proposed a novel semi-supervised strategy to optimize the sensor deployment that considered that some leak locations are unknown, and the results showed that the addition of the fuzzy c-means clustering method integrated with the original semi-JMI (joint mutual information) algorithm could efficiently improve the accuracy and stability of leak localization. Mengke Zhao et al. (2020) [26] proposed a net cost indicator based on the single-objective optimization and cost-benefit analysis models to determine an optimal set of sensor locations that would maximize the detection coverage rate under a fixed number of sensors. The method was demonstrated to be effective on a benchmark network Net3, and the sensor accuracy and pipe burst flow magnitude were shown to be key uncertainties. Narges Taravatrouy et al. (2020) [27] introduced a novel hybrid entropy-clustering approach where the redundant information from pressure sensors was reduced based on an information theory after the potential pressure sensors in WDNs were classified using a K-means clustering algorithm. Furthermore, the uncertainty of the pressure sensors' error thresholds was analyzed by a set of fuzzy members using a triangular membership. These studies have all been based on multiple leakage events, and they separately considered some practical factors. However, there is little research that comprehensively contains realistic uncertainties that can efficiently be unitized in reality. Moreover, the leakage events in the representative studies mentioned above were all generated by EPANET [28,29]—famous water system simulation software—in which a burst pipe is simulated by adding an extra demand at a target node [30].

This paper puts forward a methodology for pressure sensor layout optimization with realistic issues and considers economic budgets (reflected by sensor accuracies and sensor amounts), available sensor installation locations, sensor faults, water demand profiles, etc. The least number of pressure sensors and the layouts can firstly be gained via sequential subtraction, and the abundant sensors can then be added to compensate for the lost information caused by sensor faults on the promise of an economic budget. Furthermore, the hydraulic model [31–33] utilized in this paper to generate large number of leak samples is a novel method for simulating leakage incidents can be set at any location in WDNs rather than at only nodes. The method is applied to a representative Net3 and the results show that it has an outstanding performance with respect to applicability and practicality.

2. Mathematical model and numerical solution

In order to accurately analyze the leakage behavior of a pipeline, it is necessary to establish an advanced hydraulic model. A hydraulic model can be comprised of two aspects: one for a straight pipe, and another for a pipe network.

2.1. Straight Pipe

For analyzing the hydraulic behavior of water in a straight pipe, there are already several conventional models available. However, few of them consider water leakage in the middle of a pipe. In this paper, a mathematical model is introduced which can consider the influence of multiple leakage points.

The diameter of a pipeline is negligible compared to its length, which is generally hundreds to thousands of meters. Hence, only axial-direction water flow is considered in the proposed model. A control volume in a straight pipe (with a dip angle $\pm \theta$) with a length of dx was selected, shown in Figure 1. Time-dependent ordinary differential equations are derived from the mass and energy balances for each control volume, as shown in Figure 2. Taking the leakage loss out of a pipe wall, the mass conservation equations can be written as follows:

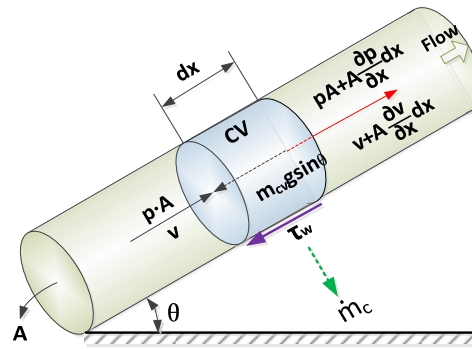


Figure 1. Water flow in a pipe.

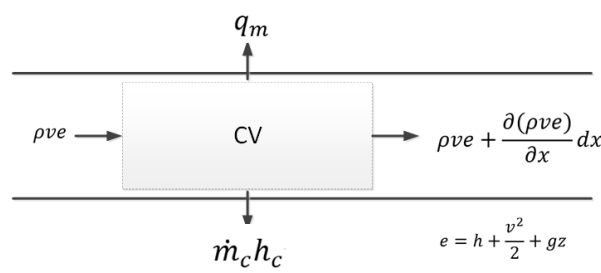


Figure 2. Energy into and out of the control volume.

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial x} = \dot{m}_c \quad (1)$$

where ρ indicates the density of the water (kg/m^3), v indicates the velocity (m/s), t indicates the time (s), x indicates the length of the pipeline (m), and \dot{m}_c indicates the specific mass loss of the water leakage ($\text{kg}/(\text{m}^3 \cdot \text{s})$).

The momentum conservation equation can be written as

$$\frac{\partial(\rho v)}{\partial t} + \frac{\partial(\rho v^2)}{\partial x} = -\frac{\partial p}{\partial x} - \frac{f \rho}{2d} v |v| - \rho g \sin \theta \quad (2)$$

The energy conservation equation can be written as

$$\frac{\partial}{\partial t} \left[\rho \left(u + \frac{v^2}{2} + gz \right) \right] + \frac{\partial}{\partial x} \left[\rho v \left(h + \frac{v^2}{2} + gz \right) \right] = \dot{m}_c \cdot h_c + q_m \quad (3)$$

where p indicates the pressure (pa), u indicates the internal energy of the water (J/kg), d indicates the inner diameter of the pipeline (m), g indicates the gravity acceleration (m/s^2), z indicates the relative height (m), h indicates the enthalpy value of the water (J/kg), q_m indicates the specific heat

loss from the water to the surrounding soil (W/m^3), and h_c indicates the enthalpy value of the water leakage (J/kg).

The state equation (Equation (4)) and the enthalpy equation (Equation (5)) for the water flow can be written as

$$\rho = \rho(p, T) \quad (4)$$

$$h = h(p, T) \quad (5)$$

Deductions about the mass, momentum, and energy conservation equations have been described many times in the literature. In this paper, a mathematical model is presented with additional water leakage modifications based on a single-phase flow simulator.

Consolidating Equations (1), (2), and (3) into a matrix form, we can obtain

$$\frac{\partial}{\partial t} \begin{bmatrix} \rho \\ \rho v \\ \rho \left(u + \frac{v^2}{2} + gz \right) \end{bmatrix} + \frac{\partial}{\partial x} \begin{bmatrix} \rho v \\ \rho v^2 + p \\ \rho v \left(h + \frac{v^2}{2} + gz \right) \end{bmatrix} = \begin{bmatrix} \dot{m}_c \\ -\frac{f\rho}{2d}v|v| - \rho g \sin\theta \\ \dot{m}_c \cdot h_c + q_m \end{bmatrix} \quad (6)$$

To summarize, the thermal-hydraulic model presented in this paper is more applicable than other models when dealing with practical conditions. There are two main improvements: 1) the material loss caused by the leakage added to the pipeline mass conversation equation; and 2) the enthalpy value of the leaked water added to the energy conversation equation.

For sensor location analysis in a water pipe network, numerical simulations of steady-state flows are sufficient due to the ultrafast wave speed of water. Therefore, the time term in Equation (6) can be omitted for solving the discretized governing equations. In this way, Equation (6) can be simplified as follows:

$$\frac{\partial}{\partial x} \begin{bmatrix} \rho v \\ \rho v^2 + p \\ \rho v \left(h + \frac{v^2}{2} + gz \right) \end{bmatrix} = \begin{bmatrix} \dot{m}_c \\ -\frac{f\rho}{2d}v|v| - \rho g \sin\theta \\ \dot{m}_c \cdot h_c + q_m \end{bmatrix} \quad (7)$$

There have been a number of developed methods that derive the discretized forms of mathematics equations. Here, the FVM (finite volume method) based on the QUICK scheme is adopted to perform spatial discretization. PISO (Pressure Implicit Split Operator) is used to solve the discrete equation. The derivation method has been detailed in a previous study [31]. Regarding this paper's focus on the WDNs, the energy conservation equation can be neglected during the numerical solution process.

2.2 Pipe Network

On basis of the graph theory, the hydraulic regime of a pipe network can be well-described. According to Kirchhoff's current and voltage laws, a pipe network's branch flow, pressure drop, and hydraulic resistance can be drawn using an analogy with the branch current, voltage, and electric resistance. In the network topology, k is the branch number and $(n+1)$ is the node number. The associated matrix A , an $n \times k$ order matrix, and the basic circuit matrix B , a $(k-n) \times k$ order matrix, of the pipe network can be gained. In the light of Kirchhoff's current law, the equation can be expressed as

$$A \cdot \dot{V}_b = 0 \quad (8)$$

where \dot{V}_b is the flow rate column vector of each branch in the pipe network $([\dot{V}_{b,1}, \dot{V}_{b,2}, \dots, \dot{V}_{b,k}])$.

From Kirchhoff's voltage law, the following can be obtained:

$$B \cdot (R_b - \rho g H_p) = 0 \quad (9)$$

where \mathbf{R}_b is the resistance column vector of each branch in the pipe network $([R_{b,1}, R_{b,2}, \dots, R_{b,k}]^T)$, \mathbf{H}_p is the pump head column vector of each branch in the pipe network $([H_{p,1}, H_{p,2}, \dots, H_{p,k}]^T)$, and the term in \mathbf{H}_p is zero unless there is a pump installed in the corresponding branch.

The resistance of the pipe branch k can be estimated by the Darcy–Weisbach equation as follows:

$$R_{b,k} = f_k \cdot \frac{8\rho L_k (\dot{V}_{b,k})^2}{\pi^2 (d_k)^5} \quad (10)$$

where L_k is the length of the pipe branch k (m), d_k is the inner diameter of the pipe branch k (m), and the friction factor f_k is estimated by the Colebrook–White equation as follows:

$$\frac{1}{\sqrt{f_k}} = -2 \lg \left(\frac{\varepsilon_k / d_k}{3.76} + \frac{2.51}{Re \sqrt{f_k}} \right) \quad (11)$$

where ε_k is the roughness of inner surface of the pipe branch k and Re is a Reynold number of the flow in the pipe branch k .

The Newton–Raphson method is typically utilized to analyze the flow rate in each branch using the nonlinear equations mentioned above. This advanced approach has been used in numerous literary works. Since the temperature of a water distribution network is relatively stable and heat loss is rare, which is different from heating a pipe network, the thermal model can be taken out of consideration. The proposed method can be used for a general network with tree-like or looped topology structures, and it is very convenient for programming. In this study, the simulation program has been developed using the Matlab platform.

3. Sensor Deployment Method

Two indexes are defined to evaluate the various pressure sensor combinations for leak identification in a WDN. One is DCR , the detection coverage rate, which ranges from zero to one and related to false negatives. The other is TDS , the total detection sensitivity, linked to false positives. A larger value in the metering data makes leakage events more obvious. If the DCR s for several pressure sensor combinations are identical, then the TDS can be utilized to identify the best option. The DCR is calculated using Equation (12) and Equation (13). If event j can be detected by a certain deployed pressure sensor within N sensors, this means that the event j is detectable, and the total number of detectable events is the $NS_{detected}$ under the corresponding N sensors. The identification of whether the j -th event E_j can be detected by the i -th sensor S_i is used to compare the pressure residual and the detection threshold at the i -th sensor node for leakage event j . The pressure residual $\Delta P_{i,j}$ is calculated by Equation (14), which can be obtained from the simulation results. The detection threshold $\Delta P_{i,j}^{threshold}$ is related to the accuracy of S_i , and in Equation (15), the normal pressure difference $\Delta P_{i,j}^{normal}$ is generally be assumed to be zero [26]. Therefore, $\Delta P_{i,j}^{threshold}$ is equal to half of the sensor accuracy δ_{sensor} . The TDS is defined by Equation (16), and the quantified the pressure residual is shown as the sensor deviation as follows:

$$DCR = \frac{NS_{detected}}{NS_{total}} = \frac{\sum_{i=1}^N \sum_{j=1}^{NS_{total}} D_{i,k}}{NS_{total}} \quad (12)$$

where the DCR is the detection coverage rate, which ranges from zero to one, $NS_{detected}$ is the total number of detectable events by the set of sensors, NS_{total} is the total number of pipe leakage events, i is the sensor index, and j is the event index, which is calculated as follows:

$$D_{i,k} = \begin{cases} 1 & \text{if } (E_j \text{ is detected by } S_i \text{ but not by } \{S_1, \dots, S_{i-1}\}) \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

where E_j is the j -th event and S_i is the i -th sensor. If $\Delta P_{i,j} \geq \Delta P_{i,j}^{threshold}$, then E_j is detected by S_i ; otherwise, it is not detected. $\Delta P_{i,j}$ is calculated as follows:

$$\Delta P_{i,j} = P_{i,j}^L - P_{i,j}^C \quad i = 1, 2, \dots, N, \quad (14)$$

where $\Delta P_{i,j}$ is the pressure residual at the i -th sensor node for leakage event j , $P_{i,j}^L$ is the pressure at the i -th sensor when leakage event j occurs, $P_{i,j}^C$ is the corresponding normal pressure at the i -th sensor under the same usage patterns of event j , and N is the number of pressure sensors. $\Delta P_{i,j}^{threshold}$ is calculated as follows:

$$\Delta P_{i,j}^{threshold} = 1/2\delta_{sensor} + |\Delta P_{i,j}^{normal}| \quad (15)$$

where $\Delta P_{i,j}^{threshold}$ is the detection threshold at the i -th sensor node for leakage event j , δ_{sensor} is the sensor accuracy, and $\Delta P_{i,j}^{normal}$ is the normal pressure difference at the i -th sensor node for leakage event j . The TDS is calculated as follows:

$$TDS = \sum_{i=1}^N \sum_{j=1}^{N_{total}} DS_{i,j} = \sum_{i=1}^N \sum_{j=1}^{N_{total}} \text{round}\left(\frac{\Delta P_{i,j}}{\delta_{sensor}}\right) \quad (16)$$

where TDS is the total detection sensitivity, which indicates the detection difficulty, where a larger value means that the events are more obvious on the devices, and $DS_{i,j}$ is the detection sensitivity of the S_i under the leakage event E_j .

The methodology consists of three steps, as shown as Figure 3. In step one, a number of leakage events are generated, which are essential for the leak identification analysis, and for this, a simulation is the most convenient channel. The simulation model employed in this paper is interpreted in part two, which can be available for a burst pipe at any location. Firstly, a hydraulic model for the WDN is established. Secondly, the schemes for the leakage events are designed and specified, including the leakage locations, burst flows, boundary conditions for the sources, tanks, users, etc. Thirdly, the simulation results for each leakage scheme and the corresponding results for a scheme with no leakage are achieved, and then the pressure divergence matrices for the N_{max} nodes can be obtained by comparing the results for leakage and no leakage events. Finally, the DCR is calculated and recognized as DCR_{max} .

In step two, the least amount and distribution of the pressure sensors are obtained through subtraction. This step is aimed at identifying the least essential pressure sensors on the promise of no reductions in the number of detection events. Beginning at the maximum value of N_{max} and deleting one from the current sensors S_1, S_2, \dots , and S_N , respectively, we can compare the maximum values for the DCR using these combinations. If the maximum DCR is equal to DCR_{max} , we can determine the largest TDS of the sensor combinations with maximum DCR and retain the corresponding $N-1$ sensors. We then repeat the previous steps to delete further sensors one by one until the maximum DCR is less than the value of DCR_{max} and the lowest number and the layout of the pressure sensors are achieved.

As for step three, the compensated pressure sensors are added as allowed by the economic budget. If N_{least} is more than N_{goal} , this indicates that the detected coverage rate has to be decreased. Under these circumstances, to subtract one from the current sensors S_1, S_2, \dots , and S_N , respectively, we retain the combination with the maximum DCR . If there is more than one status that shares the same maximum DCR , we choose the combination with the maximum TDS . We then repeat the these steps until the number of pressure sensors is decreased to N_{goal} . If N_{least} is less than N_{goal} , this means there are extra sensors can be installed in addition to those in the essential locations, and the extra sensors are determined based on the estimating the parts of the essential sensors that are broken. By subtracting one sensor included in essential sensors, as selected in step two via the Monte Carlo method, the DCR is deduced, and then we add one from the deleted sensors S_1, S_2, \dots , and $S_{N_{max}-N+1}$, respectively. The added sensor that results in the largest DCR is preserved, and if the largest DCR corresponds to more than one combination, the TDS is employed to choose the relative optimal one. Then, supposing another sensor is broken, we select the sensor that can best take the place of the broken one. These steps are repeated until the number of the least essential sensors plus the added sensors is equal to the expected value of N_{goal} .

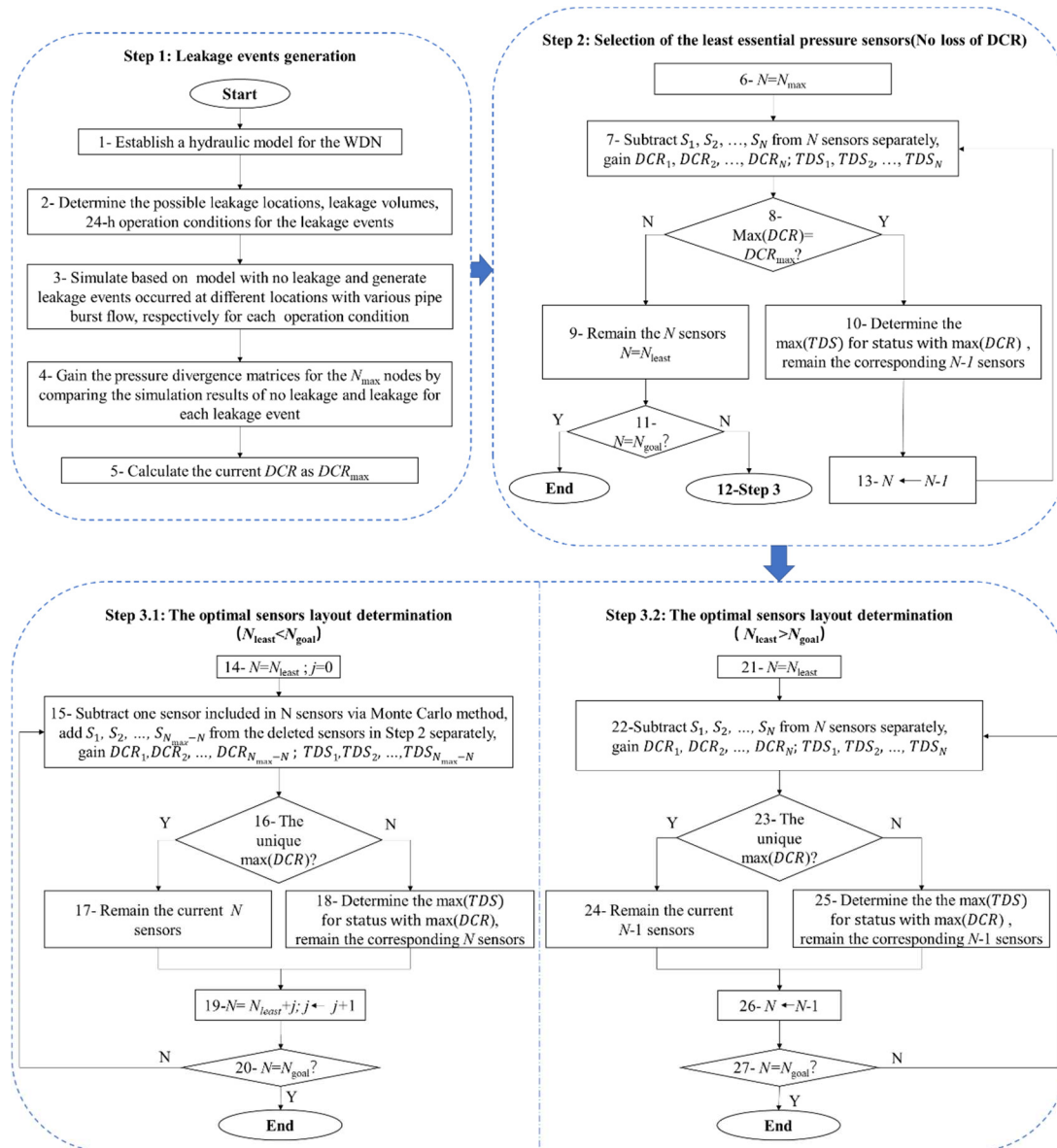


Figure 3. Flowchart of the proposed approach for pressure sensor deployment optimization for leak identification in WDNs.

4. Case study

The proposed three-stage optimization decision-making method for sensor placement was applied to a well-known example network Net3, as shown in Figure 4, the specific Net3 network specification files are publicly available and included with the EPANET Programmers toolkit [34]. Many studies have been conducted on this virtual network, including those of Schwetschenau et al. (2019) [35], Mengke Zhao et al. (2020) [26], Juan Li et al. (2019) [25], and Kegong Diao et al. (2016) [36], since real-world networks are not readily available due to security issues. Our model consists of 92 junctions, 117 pipes, 2 reservoirs (one lake and one river), and 3 tanks. To imitate the realistic fact that only partial nodes meet the objective requirements of sensor installation, 42 nodes are selected randomly to be powered and marked, as shown in Figure 2, as the potential locations where the pressure sensors can be installed. Different sensor thresholds may change the final results, and this paper is aimed at the method of instruction, without discussing the influence of the sensor thresholds. δ_{sensor} is regarded as 1 kpa. The number of expected sensors, N_{goal} , is assumed to be 21, which is equal to half of the potential nodes.

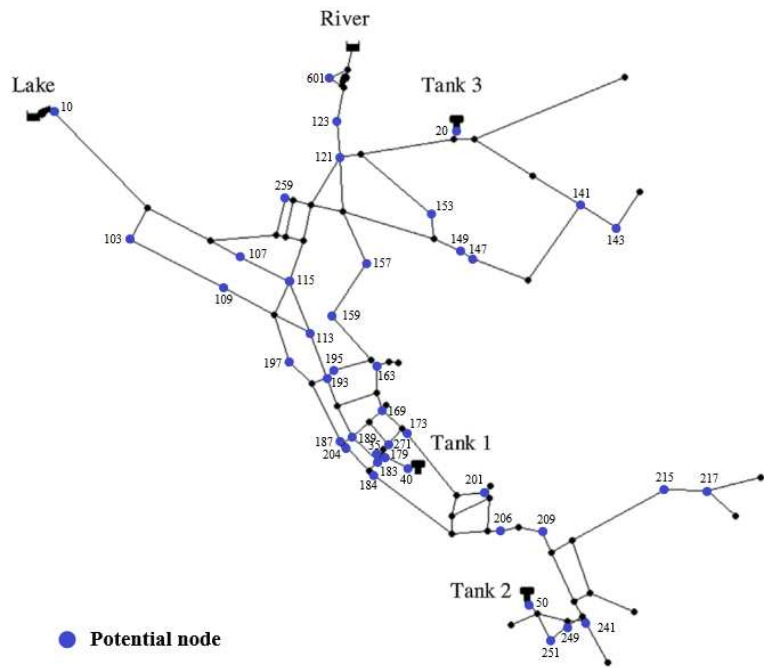


Figure 4. Layout of the Net3 network.

4.1. Leakage event generation

Using the Net3 network, a number of leakage events are generated via simulation, and we can suppose the following: (1) all pipes are considered, and the three points that are evenly distributed in each pipeline are supposed to burst; (2) all nodes are included while the reservoirs and tanks are excluded; (3) for each leakage point, the burst flow is divided into three different statuses, that is, 1%, 3%, and 6% of the flow rate under the fault-free operating condition, respectively; (4) 24 h leakage events are generated and the hourly water demand profile is set to refer to real water bill data from a certain area of China, as shown as Table 1, where the peak load is approximately 3000m³/h at 19:00 hours; and (5) the boundary condition for the reservoirs and tanks is simulated as a constant pressure for each leakage event. Therefore, the number of generated leakage events is 31896.

The D_i and DS_i for each potential node are analyzed in Figure 5, indicting the leak identification ability when there is only one pressure sensor. It can be seen that the largest D_i occurs at node 241, where the value of D_i is 19362 and that of the DCR is 60.7%, and node 601 has a superior DS_i of 56256. However, the values of both D_i and DS_i for nodes 20, 40, and 50 are zero, meaning that all leakage events have rare impacts on these nodes. These nodes are close to the three tanks, and the tanks are set at a constant pressure during the simulation, with the pressures for the tanks being relatively low compared to the supply pressures of the sources.

Table 1. Hourly water demand load rates.

Time	α	Time	α	Time	α	Time	α
1	0.0865	7	0.9312	13	0.3230	19	1.0000
2	0.0661	8	0.8195	14	0.2314	20	0.8984
3	0.0350	9	0.7670	15	0.2200	21	0.8242
4	0.0338	10	0.7626	16	0.2715	22	0.3031
5	0.1631	11	0.8355	17	0.7152	23	0.1333
6	0.7099	12	0.7764	18	0.8978	24	0.1044

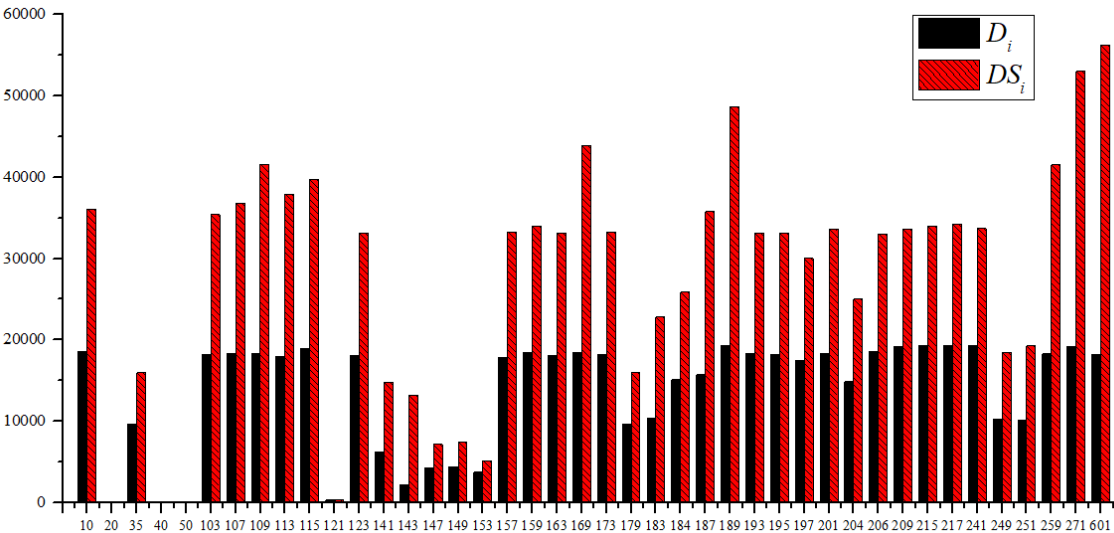


Figure 5. The D_i and DS_i values for each potential node

4.2. The least essential pressure sensor selection

In terms of the pressure sensor deployment for the 42 potential nodes, the DCR_{max} is 91.35%, and none of these 42 sensors can detect 8.65% of the leakage events mainly due to the tiny leakage volumes and marginal leak locations. Beginning at the maximum number of sensors, one sensor is removed from the 42 sensors, respectively, and the results show that when the sensor installed in node 50 is removed, the DCR is identical to the DCR_{max} and the TDS is the maximum value for this status for the DCR_{max} . Node 50 is specified to be subtracted, and then one sensor is removed from the remaining 41 sensors, respectively, and the previous steps are repeated. The progress is listed in Table 2, and it is obvious that when the number of sensors decreased from 17 to 16, the optimal DCR for the 16 sensors is less than 91.35%, indicating that some leakage events were not detected. Therefore, the minimum number of essential pressure sensors, N_{least} , is 17, and the pressure sensor layout is a combination of nodes 10, 107, 109, 113, 141, 149, 159, 169, 184, 187, 189, 195, 217, 251, 259, 271, and 601, as shown in Figure 6.

Table 2. The progress of the least essential pressure sensor selection for the Net3 network.

Number of sensors	Optimal DCR (%)	Optimal TDS	Removed location	Number of sensors	Optimal DCR (%)	Optimal TDS	Removed location
41	91.35	1163630	50	28	91.35	976387	206
40	91.35	1163630	40	27	91.35	943294	123
39	91.35	1163630	20	26	91.35	910197	193
38	91.35	1163311	121	25	91.35	877095	163
37	91.35	1158150	153	24	91.35	843807	173
36	91.35	1150979	147	23	91.35	810495	157
35	91.35	1137732	143	22	91.35	776897	209
34	91.35	1121772	35	21	91.35	743264	201
33	91.35	1105780	179	20	91.35	709547	241
32	91.35	1087334	249	19	91.35	675538	215
31	91.35	1064518	183	18	91.35	640128	103
30	91.35	1039493	204	17	91.35	600396	115
29	91.35	1009458	197	16	91.34	556527	169

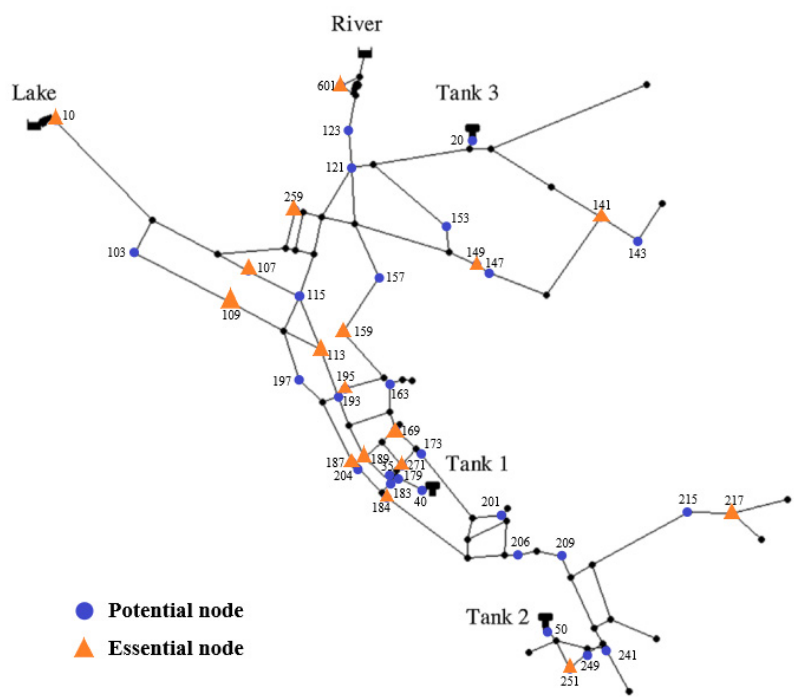


Figure 6. Layout of the least essential pressure sensors for the Net3 network.

4.3. The optimal sensor layout determination

The number of the least essential pressure sensors is 17, which is 4 less than the expected 21. There are various faults that can occur with the sensors, such as becoming stuck, damaged, etc., which leads to instrument failure. The extra four sensors are expected to take the place of the failed sensors. The progress in this section is exactly as shown in Table 3. Firstly, the Monte Carlo method is utilized and the sensor located in node 217 is supposed to have failed, which is obvious as the *DCR* declines by 2.94% and some leakage events cannot be detected. Then, 1 of the 25 sensors removed in step two is sequentially added, and node 215 is identified to have the pressure sensor installed for the best compensatory performance with a *DCR* of 91.35%. We repeat these steps and the sensor located in node 189 is removed, and nodes 197, 204, and 183 share the same *DCR* whether or not one of them takes the place of the node 189. Then, node 197 is left due to having the highest *TDS* value. Similarly, node 201 replaces node 271 and node 103 replaces node 10. Hence, the optimal 21 sensors are located in nodes 10, 107, 109, 113, 141, 149, 159, 169, 184, 187, 189, 195, 217, 251, 259, 271, 601, 215, 197, 201, and 103, respectively, as shown as Figure 7.

Table 3. The progress of the optimal sensor layout determination for the Net3 network.

Item	Current pressure sensor deployment	Indexes	
		DCR	TDS
First	I 10 107 109 113 141 149 159 169 184 187 189 195 217 251 259 271 601	91.35 %	60039 6
	II 10 107 109 113 141 149 159 169 184 187 189 195 251 259 271 601	88.41 %	56498 6
	III 10 107 109 113 141 149 159 169 184 187 189 195 215 251 259 271 601	91.35 %	59899 5
Second	I 10 107 109 113 141 149 159 169 184 187 189 195 215 251 259 271 601	91.35 %	59899 5

To evaluate the performance of the ultimate determinate sensor layout, a set of test data is generated by simulation, with no involvement in the sensor selection process. The boundary conditions, used to obtain the leakage samples in Step 1, should reflect the design scenario or typical operating condition. This ensures congruity between the load distribution of the samples and the prevailing empirical distribution patterns. Therefore, the water load of each user experiences random

fluctuations within a range of $\pm 10\%$ based on the training samples, while maintaining the other settings and topology of WDNs unaltered, generating new 24 h leakage events for test. Applying the final 21 pressure sensors to the test data, it is found that the DCR of final sensor topology is measured at 92.03% while the DCR_{\max} of 42 points stands at 92.05%. The TDS of final sensor topology is calculated as 786973 while the TDS of 42 points is 1176820. The result shows that this sensor layout exhibits a commendable leak monitoring efficacy, particularly when confronted with analogous water consumption patterns.

4.4. Impact of leakage model parameter on results

The previous section presented the application steps of the proposed method, leading to the final optimized layout. The generation of leakage samples during the initial step holds implications for the resultant outcomes. Ideally, a sample containing all possible leak locations and leak flow rates would be considered comprehensive and advantageous. However, due to computational power and time limitations, it is impracticable to comprehensively encompass all possible leakage events. Therefore, the influence of leak model parameter settings on the results is discussed in this section, to provide a reference for sample generation progress.

In section 4.1, the leakage locations are configured to encompass all nodes, as well as 1/4, 1/2, and 3/4 positions along the pipelines. The burst flow is divided into three distinct levels, namely 1%, 3%, and 6% of the flow rate calculated under fault-free operating conditions. These samples were individually selected to derive 21 optimized sensor layouts using the proposed method in this paper. Subsequently, these resulting sensor layouts are applied to the entire sample set and their corresponding DCR and TDS are determined.

The specific outcomes are shown in Table 4, with nodes arranged in reverse order of D_i .

Table 4. Optimal sensor layout obtained from different leakage samples.

Node	Total samples	Leakage location			Brust flow		
		Loc_0.25	Loc_0.5	Loc_0.75	Flow_1%	Flow_3%	Flow_6%
217	○	○	○	○	○	○	○
215	△	-	-	-	-	-	△
241	-	-	-	△	△	△	△
189	○	△	○	○	△	○	○
271	○	○	○	○	○	○	○
209	-	○	○	-	-	-	○
115	-	△	○	-	○	-	△
10	○	-	-	○	○	○	-
206	-	-	-	△	△	△	-
159	○	○	○	○	○	○	○
193	-	-	-	-	-	-	-
201	△	△	△	○	○	△	△
109	○	○	○	△	△	○	○
169	○	○	△	○	△	○	○
173	-	-	-	-	-	-	-
107	○	○	○	○	○	○	○
103	△	○	△	△	○	△	△
195	○	○	○	○	○	○	○
123	-	-	-	-	-	-	-
259	○	○	○	○	○	○	○
163	-	-	-	-	○	○	-
601	○	○	○	○	○	○	○
113	○	○	○	○	○	○	○
157	-	-	-	-	-	-	-
197	△	△	△	△	△	△	△

187	○	△	○	-	-	○	-
184	○	○	○	○	○	○	○
204	-	-	-	-	-	-	-
183	-	-	-	-	-	-	-
249	-	-	-	-	-	-	-
251	○	○	○	○	-	-	○
35	-	-	-	-	-	-	-
179	-	-	-	-	-	-	-
141	○	○	○	○	-	-	-
149	○	○	○	○	○	○	○
147	-	-	-	-	-	-	-
153	-	-	-	-	-	-	-
143	-	-	-	-	-	-	-
121	-	-	-	-	-	-	-
20	-	-	-	-	-	-	-
40	-	-	-	-	-	-	-
50	-	-	-	-	-	-	-
DCR	91.35%	91.31%	91.31%	91.32%	91.23%	91.24%	91.11%
TDS	733483	736700	736700	730479	769253	765304	753873

*○ represents the essential sensors and △ represents the supplementary sensors.

It can be seen from the table, the optimal sensor points obtained from these seven leakage samples are relatively concentrated. There are 16 nodes, such as nodes 217,189,271, etc., chosen to be final point by all seven cases. It indicates that 16 out of 21 points are identical among these results and the other 5 points are distributed from 10 nodes, such as nodes 215,241,209, etc. The leak identification abilities of the results gain by the sensor deployment method proposed are not bad with the worst *DCR* measuring at 91.11%, which means losing the ability to detect the 77 leakage events. The *DCRs* pertaining to different leak locations exhibit relative similarity and are close to the maximum *DCR* compared to the *DCRs* gained from three statuses of burst flow. This can be attributed to the limited influence of leakage positions among the identical pipeline on the pressure distribution, particularly when the pipeline is not long, compared to the magnitude of the leakage. The *TDSs* gained by separated burst flow samples are larger than others, because the chosen nodes process high detection sensitivity with decreased detection coverage rate by combination. In general, results based on a set of more comprehensive leakage samples exhibit an improved ability to reflect leakage events. Given limited computing power and time constraints, the leakage location can be set simply while placing greater emphasis on the burst flow setting. It is suggested to generate a comprehensive range of leakage events with detailed consideration of varying possible magnitudes of leakage.

5. Conclusion

This paper is aimed at pressure sensor deployment optimization in terms of leak identification in WDNs, and a methodology is proposed to achieve the best leakage fault detection performance allowed given several uncertain restrictions. The methodology is made up of three steps: Firstly, a hydraulic model is established and a number of leakage events are generated. Secondly, the least essential pressure sensors are selected by subtracting the sensors sequentially from all the potential installed nodes, with no losses in the detected events. Lastly, if the least number is less than expected, we suppose a certain sensor has failed and identify another location to best take its place within the removed sensors from the previous step, and this is repeated until the sensor number is up to the expected number and the optimal deployment is achieved. Otherwise, we subtract the sensors with the optimal *DCR* values from the expected number. A Net3 network is employed as a case study to demonstrate the specific progress of the proposed methodology, and the results show that the methodology is convenient and practical for leak identification in WDNs. The main conclusions are summarized as follows:

(1) A novel methodology is proposed in this paper to optimize the deployment of pressure sensors for leak identification in WDNs. The methodology considers economic budgets, sensor faults, and the available local conditions for installation, which are proper and practical elements for realistic projects. The economic budgets are associated with the number and accuracy of the sensors, and both of these are considered in the methodology and influence the results.

(2) Two indexes are defined to evaluate the effectiveness of sensor layout for leak identification. One is *DCR*, related to false negatives, the larger the value, the more leakage events can be reflected in the measured values. The other is *TDS*, linked to false positives, the larger the value, the higher the sensitivity of the instrument value to monitoring leakage events. The principle is to determine priority to ensure an optimal *DCR* and retain the largest *TDS* with an identical *DCR*.

(3) The lowest number of pressure sensors can be obtained via the proposed approach without leakage event detection loss, and it can be beneficial for planning developments and avoiding wasted manpower and money. Otherwise, if a budget allows, the final deployment can be obtained via Monte Carlo method by adding the compensated sensors based on the hypothesis that there has been a partial sensor failure.

(4) A well-known WDN is employed to illustrate the methodology in detail, and up to 91.35% of leakage events can be detected under the restriction of the installation locations for sensors. The results show that at least 17 pressure sensors are required to guarantee the detection of leaks, and the sensitivity of the *TDS* is approximately half of the 42 locations installed. For the optimal number of sensors (21) with a budget, the *DCR* can be 90.83%, and the *TDS* is slightly decreased when 4 sensors fail and another 4 locations are replaced.

Author Contributions: All authors contributed equally to the preparation of this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The authors sincerely acknowledge the support of the National Key Research and Development Program of China (project no. 2021YFE0116200).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Moasheri R, Jalili-Ghazizadeh M. Locating of Probabilistic Leakage Areas in Water Distribution Networks by a Calibration Method Using the Imperialist Competitive Algorithm[J]. Water Resources Management, 2020, 34(1): 35-49.
2. National Bureau of statistics of the people's Republic of China. China Statistical Yearbook [M]. Beijing: China Statistics Press, 2021 (In chinese).
3. Laspidou C S . ICT and stakeholder participation for improved urban water management in the cities of the future[J]. Water Utility Journal, 2014, 8: 79-85.
4. Kowalska B, Suchorab P, Kowalski D. Division of district metered areas (DMAs) in a part of water supply network using WaterGEMS (Bentley) software: a case study[J]. Applied Water Science, 2022, 12(7): 166 (10 pp.).
5. El-Zahab S , Zayed T . Leak detection in water distribution networks: an introductory overview[J]. Smart Water, 2019, 4(1):23-28.
6. Abdulshaheed A, Mustapha F, Ghavamian A. A pressure-based method for monitoring leaks in a pipe distribution system: A Review[J]. Renewable and Sustainable Energy Reviews, 2017, 69: 902-911.
7. Bykerk L, Valls Miro J. Detection of water leaks in suburban distribution mains with lift and shift vibro-acoustic sensors[J]. Vibration, 2022, 5(2): 370-82.
8. Fletcher R, Chandrasekaran M. Smartball-A new approach in pipeline leak detection[C]. 2008 ASME International Pipeline Conference, IPC 2008, September 29, 2008 - October 3, 2008, 2009: 117-133.
9. Lijian Y, Ying G, Songwei G. Multi-leak detection in pipeline based on optical fiber detection[J]. Optik, 2020, 220: 74-83.
10. Thusyanthan I, Blower T, Cleverly W. Innovative uses of thermal imaging in civil engineering[J]. Proceedings of the Institution of Civil Engineers, 2017, 170(2): 81-87.
11. Farooq O, Singh P, Hedabou M, et al. Machine Learning Analytic-Based Two-Staged Data Management Framework for Internet of Things[J]. Sensors, 2023: 2427 (29 pp.).
12. Keramat A, Xun W, Louati M, et al. Objective functions for transient-based pipeline leakage detection in a noisy environment: least square and matched-filter[J]. Journal of Water Resources Planning and Management, 2019, 145(10): 04019042 (13 pp.).

13. Hu Z, Tan D, Chen B, et al. Review of model-based and data-driven approaches for leak detection and location in water distribution systems[J]. *Water Supply*, 2021, 21(7): 3282-3306.
14. Creaco E, Pezzinga G. Embedding linear programming in multi objective genetic algorithms for reducing the size of the search space with application to leakage minimization in water distribution networks[J]. *Environmental Modelling & Software*, 2015, 69: 308-318.
15. Qingzhou Z, Zheng Yi W, Ming Z, et al. Leakage zone identification in large-scale water distribution systems using multiclass support vector machines[J]. *Journal of Water Resources Planning and Management*, 2016, 142(11): 04016042 (15 pp.).
16. Xuan H, Yongming H, Bin Y, et al. Novel leakage detection and water loss management of urban water supply network using multiscale neural networks[J]. *Journal of Cleaner Production*, 2021, 278: 1518-26.
17. Zhang C, Alexander B J, Stephens M L, et al. A convolutional neural network for pipe crack and leak detection in smart water network[J]. *Structural Health Monitoring*, 2023, 22(1): 232-244.
18. Daniel I, Pesantez J, Letzgus S, et al. A sequential pressure-based algorithm for data-driven leakage identification and model-based localization in water distribution networks[J]. *Journal of Water Resources Planning and Management*, 2022, 148(6): 04022025 (16 pp.).
19. Casillas A, Modera M, Pritoni M. Using non-invasive MEMS pressure sensors for measuring building envelope air leakage[J]. *Energy and Buildings*, 2021, 233: 64-75.
20. Kang D, Lansey K. Novel Approach to Detecting Pipe Bursts in Water Distribution Networks[J]. *Journal of Water Resources Planning and Management*, 2014, 140(1): 121-127.
21. Nejari F, Sarrate R, Blesa J. Optimal Pressure Sensor Placement in Water Distribution Networks Minimizing Leak Location Uncertainty[J]. *Procedia Engineering*, 2015, 119: 953-962.
22. Cheng L, Du K, Tu J P, et al. Optimal Placement of Pressure Sensors in Water Distribution System Based on Clustering Analysis of Pressure Sensitive Matrix[J]. *Procedia Engineering*, 2017, 186: 405-411.
23. Miquel . Cuguero, Puig V , Quevedo J . Optimal Pressure Sensor Placement and Assessment for Leak Location Using a Relaxed Isolation Index: Application to the Barcelona Water Network[J]. *Control Engineering Practice*, 2017, 63(C):1-12.
24. Ferreira B, Carrico N, Covas D. Optimal Number of Pressure Sensors for Real-Time Monitoring of Distribution Networks by Using the Hypervolume Indicator[J]. *Water*, 2021, 13(16): 2235 (16 pp.).
25. Li J, Wang C, Qian Z, et al. Optimal sensor placement for leak localization in water distribution networks based on a novel semi-supervised strategy[J]. *Journal of Process Control*, 2019, 82: 13-21.
26. Zhao M K, Zhang C, Liu H X, et al. Optimal sensor placement for pipe burst detection in water distribution systems using cost-benefit analysis[J]. *Journal of Hydroinformatics*, 2020, 22(3): 606-618.
27. Taravatrooy N, Nikoo M R, Hobbi S, et al. A novel hybrid entropy-clustering approach for optimal placement of pressure sensors for leakage detection in water distribution systems under uncertainty[J]. *Urban Water Journal*, 2020, 17(3): 185-98.
28. Sela L, Salomons E, Housh M. Plugin prototyping for the EPANET software[J]. *Environmental Modelling and Software*, 2019, 119: 49-56.
29. Muller L, Gericke J, Pietersen J. Methodological approach for the compilation of a water distribution network model using QGIS and EPANET[J]. *Journal of the South African Institution of Civil Engineering*, 2020, 62(4): 32-43.
30. Muranho J, Ferreira A, Sousa J, et al. Pressure-dependent Demand and Leakage Modelling with an EPANET Extension - WaterNetGen[J]. *Procedia Engineering*, 2014, 89: 632-9.
31. Hai W, Haiying W, Tong Z, et al. A novel model for steam transportation considering drainage loss in pipeline networks[J]. *Applied Energy*, 2017, 188: 178-89.
32. Hai W, Wang H, Zhou H, et al. Modeling and optimization for hydraulic performance design in multi-source district heating with fluctuating renewables[J]. *Energy Conversion and Management*, 2018, 156: 113-129.
33. Hai W, Hua M. Improved thermal transient modeling with new 3-order numerical solution for a district heating network with consideration of the pipe wall's thermal inertia[J]. *Energy*, 2018, 160: 171-83.
34. USEPA. 2020. "EPANET." Accessed May 13, 2023. <https://www.epa.gov/water-research/epanet>.
35. Schwetschenau S E, Vanbriesen J M, Cohon J L. Integrated Multiobjective Optimization and Simulation Model Applied to Drinking Water Treatment Placement in the Context of Existing Infrastructure[J]. *Journal of Water Resources Planning and Management*, 2019, 145(11): 04019048 (11 pp.).
36. Diao K, Sweetapple C, Farmani R, et al. Global resilience analysis of water distribution systems[J]. *Water Research*, 2016, 106: 383-393.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.