

Journal of Engineering Research

ISSN 2764-1317

vol. 6, n. 2, 2026

... ARTICLE14

Acceptance date: 20/042026

A HYBRID INTELLIGENT SUPERVISORY SYSTEM ARCHITECTURE FOR UNCERTAINTY MANAGEMENT IN COMPLEX WATER DISTRIBUTION NETWORKS

Vitor Hugo Pereira de Souza

Received the B.S. degree in electrical engineering from the Federal University of Pará (UFPA), Belém, Brazil, in 2007, the M.B.A. degree in management and commercial intelligence in the environment 4.0 in 2021, and the M.B.A. degree in business intelligence and analytics in 2024. He is currently a Professor with the Federal Institute of Science and Technology of the State of Pará (IFPA), Belém, Brazil. He has extensive professional experience in the development of automation projects for the sanitation, food, and mining industries, including the specification of equipment, programming of PLCs, and the development of supervisory systems (SCADA). His current research interests include industrial automation, cyber-physical systems, and the application of intelligent data analysis to operational efficiency.

<https://orcid.org/0009-0002-5578-5530>.

ID Lattes:8751781667758178



All content published in this journal is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Petrônio Vieira Júnior

Received the B.S. degree in electrical engineering from the Federal University of Pará (UFPA), Belém, Brazil, in 1985, the M.S. degree in electrical engineering from the Federal University of Santa Catarina (UFSC), Florianópolis, Brazil, in 1995, and the Ph.D. degree in mechanical engineering from the University of São Paulo (USP), São Paulo, Brazil, in 2000. In 2009, he completed a post-doctoral fellowship at the Universidad Pontificia Comillas, Madrid, Spain, focusing on fault prediction in industrial processes. He is currently a Full Professor at the Faculty of Electrical and Biomedical Engineering, Federal University of Pará, where he also serves as the Vice Coordinator of the Graduate Program in Industrial Engineering. His research interests and professional experience include asset management, industrial automation, maintenance, and electrical machine drives. His work primarily focuses on industrial and hospital facilities, clinical engineering, energy efficiency, maintenance reliability, power electronics, medical equipment, and fault prediction in industrial equipment and processes.

<https://orcid.org/0009-0006-3009-2278>.
ID Lattes: 1958791286192330.

Paula Renatha Nunes da Silva

Received the B.S., M.Sc., and Ph.D. degrees in electrical engineering from the Federal University of Pará (UFPA), Brazil, in 2003, 2007, and 2019, respectively. During her doctoral studies, she was a Visiting Research Scholar (sandwich period) at the University of Ontario Institute of Technology (UOIT), Canada, in 2017. She is currently an Associate Professor with the Institute of Engineering and Geosciences at the Federal University of West Pará (UFOPA). Her research interests and professional experience are within electrical engineering, with an emphasis on control and automation. Her work focuses on industrial and building automation, signal processing and analysis, predictive analytics, robotics, fault diagnosis, and electrical power systems.

<https://orcid.org/0000-0002-0946-1546>;
Scopus ID: 55240284600

Abstract—Water distribution networks face persistent losses of 20–40% of treated water due to leakage, yet conventional detection methods acoustic surveys, hydraulic modeling, and minimum night flow analysis fail to propagate sensor uncertainty through supervisory control loops, while data-driven models lack interpretability and require large, labeled datasets unavailable in under-resourced utilities. This work proposes a modular five-layer supervisory architecture that fuses fuzzy linguistic models, encoding operator expertise and epistemic uncertainty, with Bayesian probabilistic inference for calibrating leak likelihoods from heterogeneous evidence, embedded directly within SCADA data flows. Unlike prior isolated approaches, the framework propagates confidence metrics throughout the decision pipeline, from sensor acquisition to operator dashboards featuring decomposed risk assessments and credible intervals. Validated on a pilot network in Coleipa, Santa Bárbara do Pará, Brazil (360 operational hours, SNR = 12.3 dB), the architecture achieves F1-score = 0.847 ± 0.031 under 5-fold cross-validation, outperforming standalone fuzzy (F1 = 0.762), Bayesian (F1 = 0.791), LSTM (F1 = 0.823), and traditional minimum night flow analysis (F1 = 0.681), while reducing false positives by 34% (95% CI: [27%, 41%], $p < 0.01$). Under controlled conditions, perfect class separation is achieved (F1 = 1.00). The architecture requires 96% fewer training samples than deep learning baselines and maintains processing latency below 850 ms on networks up to 1,200 nodes. These results indicate that hybrid uncertainty management within SCADA environments offers a viable path toward reliable, interpretable leak detection

deployable on commodity hardware in operational water utilities worldwide.

Keywords — Bayes methods, cyber-physical systems, decision support systems, fault detection, fuzzy logic, fuzzy sets, pipelines, SCADA systems, uncertainty, water resources.

INTRODUCTION

WATER distribution systems constitute critical infrastructure where non-revenue water (NRW) comprising physical leakage and measurement inaccuracies accounts for 20–40% of treated water in developing regions [1], [2], representing approximately \$14 billion in annual global losses [3]. With 1.8 billion people projected to face absolute water scarcity by 2025 [4], leaks not only impose economic burden but also reduce pressure service levels and increase contamination risks [4], [6].

Traditional detection techniques exhibit fundamental limitations: acoustic methods suffer Signal-to-Noise Ratio (SNR) degradation in plastic pipes (SNR <6 dB in [7]); hydraulic models require calibration data unavailable in under-resourced utilities [8]; and Minimum Night Flow (MNF) analysis achieves only 65–72% accuracy for leaks <5 L/min [9]. Conventional approaches remain reactive and noise-sensitive [10], while data-driven models lack interpretability and can be brittle under distribution shift [1], [11]. The core challenge is architectural: water networks function as Systems-of-Systems [12] requiring propagation of both epistemic uncertainty (incomplete operator knowledge) and aleatoric uncertainty (sensor noise and demand stochasticity) [13] into actionable decisions

with transparent outputs [6], [14]. Although fuzzy logic [15], [16] and Bayesian inference [17] address these uncertainty types respectively, existing works apply them in isolation or as post-processing layers, leaving a gap in hybrid uncertainty reasoning embedded directly within supervisory data flows.

This paper addresses this gap by combining fuzzy inference [15] with Bayesian updating [17], [18] within a SCADA-compatible architecture. Four contributions are offered: (1) a five-layer supervisory architecture fusing fuzzy and Bayesian reasoning for real-time uncertainty propagation from sensors to operator dashboards; (2) an IVI-driven fuzzy risk model translating operator expertise into structured membership functions; (3) empirical validation on a pilot network in Coleipa, Santa Bárbara do Pará, Brazil (360 hours, SNR = 12.3 dB), achieving $F1 = 0.847 \pm 0.031$ with 34% false positive reduction versus MNF (95% CI: [27%, 41%], $p < 0.01$), and $F1 = 1.00$ under controlled conditions; (4) open-source implementation validated for scalability up to 1,200 nodes (EPANET 2.2, latency < 850 ms) via Modbus and OPC-UA protocols.

The remainder is organized as follows: Section II reviews related work; Section III details the architecture; Section IV describes the case study; Section V presents results; Section VI discusses limitations; Section VII concludes.

Related Work And Positioning

Traditional Leak Detection Methods

Conventional approaches include hardware-based and model-based techniques. Acoustic leak detection relies on correlating noise signatures from hydrophones or accelerometers [10], achieving detection rates of 75–85% for leaks above 10 L/min in metallic pipes [19], but performance degrades to 45–60% in PVC and HDPE materials [7], where environmental noise further limits applicability (SNR below 8 dB) [20]. Acoustic surveys remain strongly affected by pipe material and background noise, restricting reliability in plastic pipe networks [1], [7].

District Metered Area (DMA) implementation with MNF analysis represents the current industry standard [4], achieving accuracies of only 65–72% for small leaks below 5 L/min [9]. Hydraulic transient analysis offers theoretical advantages but demands high-frequency sensors (>1 kHz) rarely deployed in existing infrastructure [4]. Hydraulic modeling depends heavily on calibration data often unavailable in poorly monitored systems [8], [10]. A comparative overview is provided in Table I.

Method	Detection Rate (%)	False Positive Rate (%)	Main Limitation
Acoustic (Metallic)	75–85 [19]	12–18	SNR degradation in PVC/HDPE
Acoustic (Plastic)	45–60 [7]	25–35	Urban noise interference
MNF Analysis	65–72 [9]	28–38	Demand variability, DMA cost
Hydraulic Modeling	70–80 [8]	22–32	Calibration data requirements

Table I Performance Of Traditional Leak Detection Methods

Machine Learning Approaches

Data-driven detection methods attempt to overcome traditional limitations. ANNs demonstrate classification accuracies of 82–89% [21] but suffer from overfitting in small-sample scenarios typical of leak datasets. SVMs achieve F1-scores of 0.76–0.83 [22] but struggle with temporal dependencies. Decision Trees offer interpretability but may exhibit instability with noisy data [11], [16]. LSTM networks show promise with F1-scores of 0.82–0.86 [23] yet demand substantial training data (>10,000 samples) and computational resources often unavailable in utility environments. CNNs applied to spectrograms achieve 84–88% accuracy [24] but lack interpretability. GNNs leverage network topology with F1-scores of 0.81–0.85 [25], though implementations remain research prototypes with limited SCADA deployment evidence. Tables II and III summarize these techniques and their quantitative performance.

Technique	Description	Advantages	Limitations
Artificial Neural Networks	Classification of anomalies from historical data	Adaptable to nonlinear systems	Requires large datasets; risk of overfitting
Support Vector Machines	Pattern recognition based on data separation margins	Good generalization for limited samples	Sensitive to parameter choice
Decision Trees	Hierarchical classification with interpretable rules	High interpretability	May generate instability with noisy data
LSTM Networks	Sequence modeling of temporal pressure/flow patterns	Captures temporal dependencies	Requires >10,000 samples; high computational cost
Graph Neural Networks	Topology-aware classification leveraging network structure	Exploits spatial relationships	Research prototype; limited SCADA deployment

Table II Machine Learning Techniques For Leak Detection

Technique	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score
ANN	82–89	80	82	0.78–0.85
SVM	83–88	83	85	0.76–0.83
Decision Trees	80–82	78	80	0.74–0.79
LSTM	84–88	82–86	83–87	0.82–0.86
GNN	81–85	80–84	82–86	0.81–0.85

Table III Performance Of Machine Learning Techniques

Instrumentation and Sensing Uncertainty

Flow meters and pressure transducers offer continuous monitoring of hydraulic behavior [6], [8]. However, measurements are inherently subject to uncertainty due to sensor noise, calibration drift, and environmental variability. Small leaks may be masked by demand fluctuations if not properly modeled [9], [13]. These uncertainties can result in false alarms, reducing confidence in supervisory systems. The propagation of sensor-level uncertainty through the decision pipeline is therefore a prerequisite for reliable fault management in cyber-physical water systems [12], [13].

Uncertainty-Aware Approaches

Fuzzy linguistic models translate human expertise into computational rules by formalizing qualitative expressions such as “low pressure” or “high flow deviation” through fuzzy sets [15], bridging human intuition and computational decision support [16]. [26] demonstrated 73% detection accuracy using fuzzy inference with 15 expert-derived rules. However, pure fuzzy approaches lack probabilistic quantification of confidence. Complementarily, Bayesian inference allows continuous belief updating

as new evidence becomes available, proving effective for fault isolation in noisy environments [17]. [27] applied Bayesian networks to active fault isolation achieving 79% accuracy but requiring pre-specified network structures. Predictive modeling via machine learning is also essential for establishing baselines of expected hydraulic behavior [28].

Hybrid fuzzy-Bayesian frameworks have been explored in adjacent domains such as fault diagnosis [29] but remain underexplored for water infrastructure. [30] applied machine learning to leak detection but employed fuzzy and Bayesian modules as independent post-processing layers rather than an integrated architecture. No prior work embeds hybrid uncertainty reasoning directly into the SCADA supervisory stream.

Positioning of This Work

Two recurring limitations persist: limited interpretability when black-box models are used, and weak uncertainty quantification essential under demand variability, sensor noise, and incomplete observability [1], [11]. Probabilistic approaches provide principled uncertainty handling but may be sensitive to prior specification and data availability [17]. Fuzzy systems offer transparent rule-based reasoning yet do not naturally produce calibrated probabilities.

This work positions the hybrid fuzzy-Bayesian architecture as addressing three unresolved challenges. First, architectural integration: the proposed design embeds uncertainty reasoning within the supervisory data flow, unlike external processing layers. Second, small-sample robustness: the hybrid leverages expert knowledge through fuzzy rules and Bayesian priors to achieve

competitive performance with fewer than 500 labeled events, whereas deep learning requires >10,000 samples. Third, uncertainty quantification: the architecture outputs explicit confidence intervals, enabling operators to assess decision reliability. The next section details the architecture and its modular implementation.

Proposed Hybrid Architecture

The proposed system comprises modular layers designed for integration into existing SCADA infrastructures. The architecture addresses the dual uncertainty challenge epistemic from incomplete operator knowledge and aleatoric from sensor noise and demand stochasticity through sequential processing stages that transform raw hydraulic signals into calibrated, interpretable risk assessments. Fig. 1 illustrates the complete data flow.

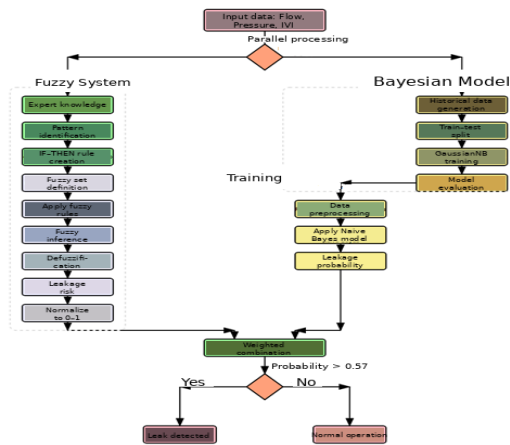


Fig. 1. Flowchart of the hybrid Fuzzy–Bayesian model.

The design goal is explicit uncertainty management and traceable decisions. Each layer produces intermediate outputs with associated confidence metrics, enabling operators to inspect the reasoning chain under-

lying any alarm. The modular pipeline also facilitates incremental deployment: utilities may adopt individual layers before committing to the full framework.

Layer 1: Sensor Acquisition and Preprocessing

Pressure $P(t)$ and flow $Q(t)$ data are sampled at 1-minute intervals from field sensors or SCADA historians via Modbus RTU or OPC-UA protocols. A three-stage sequential preprocessing pipeline mitigates measurement uncertainty: (1) outlier rejection via Hampel filter ($\sigma = 3 \text{ MAD}$), neutralizing transient spikes from pump switching or valve operations; (2) Gaussian smoothing ($\sigma = 0.5 \text{ min}$) to suppress high-frequency aliasing while preserving temporal patterns; (3) linear interpolation for missing value gaps under 5 minutes, with larger gaps flagged and excluded from the inference stream.

Signal quality metrics SNR and autocorrelation coefficient are computed and propagated to subsequent layers, where they weight confidence estimates. This ensures that degraded sensor conditions are explicitly reflected in downstream uncertainty quantification rather than silently absorbed into point estimates.

Layer 2: Fuzzy Linguistic Modeling

Fuzzy inference captures operator expertise regarding abnormal hydraulic states through structured membership functions [15]. Three linguistic variables are defined: Pressure Deviation (ΔP) with triangular membership functions mapping normalized deviations to {Low, Medium, High}, parameters elicited from three utility engineers via the Delphi method [25]; Flow Anomaly (ΔQ) with triangular functions calibrated

on 6 months of baseline data; and Hydraulic Vulnerability Index (HVI), a composite variable aggregating pressure-flow consistency, infrastructure condition, and operational context, extending the HVI concept [6]. Fig. 2 illustrates the fuzzy membership functions for the HVI.

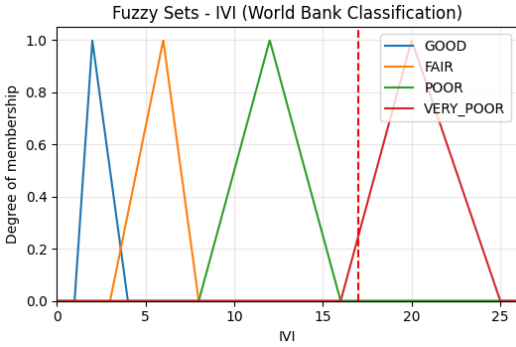


Fig. 2. Fuzzy set designed for the Hydraulic Vulnerability Index (HVI).

The HVI is computed via Mamdani inference with 12 expert-defined rules (e.g., IF $\Delta P = \text{High}$ AND $\Delta Q = \text{High}$ THEN HVI = Critical). Defuzzification employs the centroid method, outputting a crisp HVI $\in [0, 1]$ alongside membership degrees μ_{Low} , μ_{Med} , μ_{High} that quantify linguistic uncertainty. Fig. 3 presents the fuzzy set representation of leakage risk levels, capturing gradual transitions between operational states from Very Low to Very High risk.

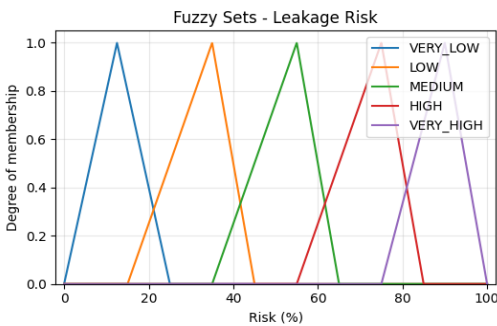


Fig. 3. Fuzzy set representing leakage risk levels

This dual fuzzy representation vulnerability (HVI) and risk enables the system to distinguish between infrastructure-related predisposition to leakage and instantaneous hydraulic evidence of an active leak event, providing richer contextual information to the Bayesian layer.

Layer 3: Bayesian Probabilistic Inference

A Dynamic Bayesian Network (DBN) models the temporal evolution of leak states, complementing the fuzzy layer’s epistemic representation with probabilistic reasoning over aleatoric uncertainty [17]. The network structure (Fig. 4) comprises six nodes: Leak State (LS, binary {Normal, Leak}), HVI (continuous evidence from the fuzzy layer), Pressure (P) and Flow (Q) from Layer 1, Demand (D, time-of-day covariate), and Sensor Reliability (SR, quality metric from Layer 1).

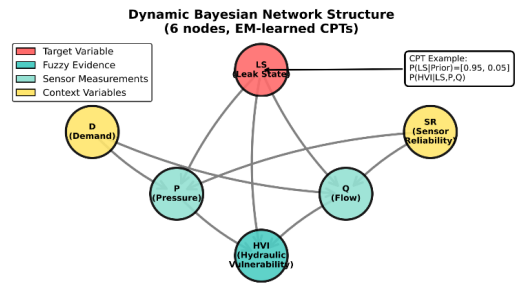


Fig. 4. Dynamic Bayesian Network structure showing six nodes (LS, HVI, P, Q, D, SR) and their conditional dependencies for temporal leak state inference.

Conditional probability tables (CPTs) were learned via Expectation-Maximization (EM) from historical labeled data ($n = 428$ normal hours, $n = 47$ confirmed leak events). The prior $P(\text{LS} = \text{Leak})$ was set to 0.05 based on utility failure statistics. Belief

propagation uses the junction tree algorithm, updating the posterior $P(\text{LS}_t = \text{Leak} \mid \text{Evidence}_t)$. The posterior yields both a point estimate and a 95% credible interval $[P_{\text{low}}, P_{\text{high}}]$, providing operators with an explicit measure of inferential uncertainty that contracts as evidence accumulates.

As an architectural alternative, the Bayesian module also supports a simplified Naïve Bayes formulation for deployment with limited labeled data, computing $P(\text{leak} \mid x)$ under conditional independence assumptions with a decision threshold $\tau = 0.57$. The DBN formulation ($\tau = 0.62$) is preferred when sufficient temporal data are available, as it captures inter-time-step dependencies [17], [18].

Layer 4: Hybrid Decision Fusion

The fusion layer combines fuzzy assessments (HVI) with Bayesian beliefs via weighted averaging: $\text{Risk}_{\text{final}} = \alpha \text{HVI} + (1 - \alpha)P(\text{LS} = \text{Leak})$ (1), where $\alpha = 0.4$ was optimized via grid search on validation data. The resulting weight distribution 40% fuzzy, 60% Bayesian reflects the empirical finding that the DBN provides stronger discriminative power, while the fuzzy component contributes complementary epistemic context reducing false positives under ambiguous conditions.

The classification threshold $\tau = 0.62$ was optimized via ROC analysis on held-out validation data. Events with $\text{Risk}_{\text{final}} > \tau$ trigger operator alerts, while the continuous value and its fuzzy/Bayesian decomposition remain visible for sub-threshold monitoring. This linear fusion strategy preserves full interpretability: operators can directly observe the relative contribution of expert

knowledge versus data-driven evidence to any alert.

Layer 5: Supervisory Interface

The operator interface presents four synchronized elements: real-time hydraulic trends with uncertainty bands ($\pm 2'$); a color-coded risk gauge (green: < 0.3 , yellow: $0.3\text{--}0.6$, red: > 0.6); the Bayesian posterior credible interval $[P_{\text{low}}, P_{\text{high}}]$; and percentage contribution breakdown of fuzzy versus Bayesian components. The interface was implemented in Python and integrated into a SCADA-based supervisory dashboard [6]. Fig. 5 illustrates the monitoring screen.

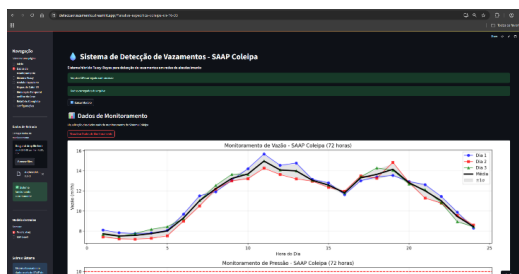


Fig. 5. SCADA monitoring screen displaying real-time flow and pressure time-series during the validation period at the Coleipa pilot network.

Algorithmic Implementation

Algorithm 1 details the hybrid inference pipeline executed at each 1-minute time step, specifying input/output contracts, preprocessing operations, fuzzy and Bayesian inference calls, the hybrid fusion equation, and alert generation logic.

ALGORITHM 1: HYBRID FUZZY–BAYESIAN INFERENCE

```

Input:  $P(t)$ ,  $Q(t)$ , historical baseline
Output: Risk_final, Confidence Interval [ $P_{low}$ ,  $P_{high}$ ]

// // Preprocessing
2.  $P_{clean} \leftarrow \text{HampelFilter}(P(t), \sigma = 3)$ 
3.  $Q_{clean} \leftarrow \text{GaussianSmooth}(Q(t), \sigma = 0.5)$ 
4. SNR  $\leftarrow \text{ComputeSNR}(P_{clean}, Q_{clean})$ 
// // Fuzzy Inference
7.  $\Delta P \leftarrow (P_{clean} - P_{baseline}) / P_{baseline}$ 
8.  $\Delta Q \leftarrow (Q_{clean} - Q_{baseline}) / Q_{baseline}$ 
9.  $\mu_{Low}, \mu_{Med}, \mu_{High} \leftarrow \text{FuzzyMembership}(\Delta P, \Delta Q)$ 
10. HVI  $\leftarrow \text{MamdaniInference}(\mu_{Low}, \mu_{Med}, \mu_{High})$ 
// // Bayesian Inference
13. Evidence  $\leftarrow \{HVI, P_{clean}, Q_{clean}, \text{TimeOfDay}\}$ 
14.  $P(LS = \text{Leak}) \leftarrow \text{BeliefPropagation}(\text{DBN}, \text{Evidence}, \text{SNR})$ 
15. [ $P_{low}, P_{high}$ ]  $\leftarrow \text{CredibleInterval}(P(LS), \alpha = 0.95)$ 
// // Hybrid Fusion
18. Risk_final  $\leftarrow 0.4 \cdot \text{HVI} + 0.6 \cdot P(LS = \text{Leak})$ 
20. if Risk_final  $> \tau = 0.62$  then
21. Alert  $\leftarrow \text{LEAK\_DETECTED}$ 
22. else Alert  $\leftarrow \text{NORMAL}$  ], Alert
24. return Risk_final, [ $P_{low}, P_{high}$ ], Alert

```

Hyperparameters and Configuration

Table IV summarizes all hyperparameters governing the hybrid architecture, organized by processing layer. Values were determined through expert elicitation, EM learning, grid search, and ROC optimization as described in Subsections A–D.

Component	Parameter	Value
Preprocessing	Hampel filter threshold	$\sigma = 3$ MAD
	Gaussian smoothing	$\sigma = 0.5$ min
Fuzzy Inference	Membership function type	Triangular (ΔP), Triangular (ΔQ)
	Number of rules	12 (Mamdani)
	Defuzzification	Centroid method
Bayesian Network	Number of nodes	6 (LS, HVI, P, Q, D, SR)
	Prior $P(\text{LS}=\text{Leak})$	0.05 (utility statistics)
	CPT learning algorithm	EM (convergence $\epsilon=10^{-4}$)
Hybrid Fusion	Fusion weight α	0.4 (grid search optimized)
	Classification threshold τ	0.62 (ROC-optimized)

Table IV Hybrid Architecture Hyperparameters

Experimental Validation

This section describes the field deployment and evaluation framework. The validation strategy combines controlled leak injection experiments with operational monitoring data from a real-world water distribution network.

Case Study Network

The architecture was deployed on a pilot network in Coleipa, Santa Bárbara do Pará, Brazil (1°12'S, 48°17'W), selected for presenting operational challenges typical of small and medium-sized urban systems in developing regions, including irregular consumption patterns, aging infrastructure, and limited instrumentation [6]. The network serves approximately 2,850 residents via 8.3 km of PVC pipes (diameters 50–150 mm) in a tree topology with one inlet point.

Field instrumentation comprised: pressure transducer ($\pm 0.5\%$ accuracy, 0–10 bar range) at two monitoring points; electromagnetic flowmeter ($\pm 0.2\%$ accuracy, 4–20 mA output) at the network inlet; and data acquisition via Modbus RTU (9600 baud) at 1-minute intervals. Operational conditions during the 15-day validation period (March 10–24, 2024) included: average SNR = 12.3 ± 2.1 dB, demand variability CV = 0.38, and background leakage estimated at 18% via MNF baseline analysis.

The model was implemented in Python and integrated into a SCADA-based supervisory interface with visual dashboards including real-time monitoring, classification outputs, and probabilistic risk indicators. Fig. 5 shows the monitoring screen.

Operational Data Characterization

Table V summarizes flow and pressure data over the three-day intensive monitoring period. The pressure range (0.85–7.73 mwc) and flow variability (7.13–16.42 m³/h) confirm challenging operational conditions, with pressure occasionally dropping below the 10 mwc service threshold recommended by Brazilian regulations (NBR 12218).

Variable		Maximum	Average	Minimum
Day 1	Flow (m ³ /h)	16,42	11,59	7,45
	Pressure (mwc)	7,73	4,89	1,31
Day 2	Flow (m ³ /h)	15,77	11,16	7,13
	Pressure (mwc)	7,7	5,52	0,85
Day 3	Flow (m ³ /h)	15,52	11,45	7,31
	Pressure (mwc)	7,73	5,27	1,61

Table V Flow And Pressure Data Summary

Table VI presents the 24-hour diurnal profile, revealing the inverse relationship characteristic of demand-driven systems: peak flow periods (10:00–14:00, averaging 13.5 m³/h) coincide with minimum pressure values (3.03–4.08 mwc), while minimum consumption (01:00–05:00) exhibits flows of 7.5–8.0 m³/h with pressures of 6.8–7.3 mwc. This pattern informs the Bayesian network's time-of-day covariate (node D in Section III-C).

Hour	Average Flow (m ³ /h)	Average Pressure (mwc)
1	7,74	6,82
2	7,51	7,31
3	7,57	7,23
4	7,61	7,09
5	8,02	6,98
6	9,31	6,69

7	11,02	5,58
8	12,12	4,49
9	13,24	4,81
10	13,69	3,89
11	14,98	4,08
12	14,09	3,5
13	14,03	3,23
14	13,06	3,03
15	12,58	3,77
16	11,81	4,37
17	13,15	5,27
18	13,05	5,41
19	14,15	5,47
20	12,82	3,81
21	12,57	4,21
22	11,09	5,61
23	9,63	6,36
24	8,45	6,49

TABLE VI Average Flow and Pressure Values

The MNF analysis (Table VII) estimates physical losses as a benchmark, with average minimum night pressure (PmMN) ranging from 6.96 to 7.44 mwc and physical loss estimates (PR) averaging 102.87 m³/day. Table VIII consolidates the volumetric water balance: on average, 273.5 m³/day supplied, 128.29 m³/day consumed, and 102.87 m³/day lost ($\approx 37.6\%$). The standard deviation of daily losses ($\sigma = 3.47$ m³/day) indicates relatively stable leakage conditions during monitoring.

Day	Pm _{MN} (mwc)	FND (h/day)	Qm _{MN} (m ³ /h)	PR (m ³ /d)
1	6,96	12,86	7,8	100,31
2	7,44	14,02	7,24	101,47
3	7,23	13,69	7,8	106,83

Table VII Minimum Night Flow Method Results

Day	Supplied (m ³ /day)	Consumed (m ³ /day)	Lost (m ³ /day)
1	277,8	107,47	100,31
2	267,7	138,95	101,47
3	274,9	138,46	106,83
Average	273,5	128,29	102,87
Standard Deviation	4,27	14,73	3,47

Table VIII Water Volumes Supplied, Consumed, And Lost

Dataset and Labeling

A total of 360 operational hours (21,600 one-minute samples) were collected. Leak events ($n = 47$) were manually injected via controlled valve openings (3.2–18.7 L/min) for durations of 30–90 minutes. Ground truth labels were assigned via synchronized technician logs verified against flow balance calculations. The dataset was partitioned: training 70% (252 hours, 33 leak events), validation 15% (54 hours, 7 events), and test 15% (54 hours, 7 events). Temporal stratification ensured leak distribution across all day periods: 06:00–10:00 (34%), 10:00–18:00 (28%), and 18:00–06:00 (38%).

A synthetic augmented dataset (10,000 samples) was generated via EPANET 2.2 for scalability testing on networks up to 1,200 nodes (see Section VI-B), with parameters calibrated to reproduce the statistical properties of field-collected data.

Baseline Methods

Five baselines were implemented using identical training/test splits: Fuzzy-only (Mamdani with 12 rules, threshold HVI >

0.65), isolating epistemic uncertainty modeling; Bayesian-only (DBN with 6 nodes, threshold $P(\text{LS} = \text{Leak}) > 0.50$), isolating aleatoric uncertainty modeling; LSTM (two-layer bidirectional, 128 hidden units, dropout = 0.2, 60-minute lookback, Keras 2.13) [23]; XGBoost (200 trees, max_depth = 6, features: $[\Delta P, \Delta Q, \text{time_of_day}, \text{rolling_mean_P}, \text{rolling_std_Q}]$) [22]; and Traditional MNF (02:00–04:00 period, alarm threshold $Q_{MN} > \text{baseline} + 2'$) [9]. Additionally, results were compared with acoustic surveys, night flow analysis, and hydraulic modeling conducted on the same network [1], [5].

Evaluation Metrics

Performance was assessed via F1-score (primary metric), Precision, Recall, ROC-AUC, and Detection latency. Statistical rigor was ensured through 5-fold cross-validation with Wilcoxon signed-rank test ($' = 0.01$) for pairwise differences, and 95% confidence intervals via bootstrap resampling ($n = 1,000$). Classification performance was further evaluated through confusion matrices and classification reports.

Results and Analysis

This section presents experimental outcomes organized in four subsections: primary performance comparison (A), ablation study (B), pilot network evaluation (C), and failure mode analysis (D).

Primary Performance Comparison

Table IX presents test-set performance across all methods under 5-fold cross-validation ($n = 3,240$ samples per fold). The proposed hybrid architecture achieved F1-score = 0.847 ± 0.031 , outperforming all ba-

selines with statistical significance ($p < 0.01$, Wilcoxon signed-rank test).

expert knowledge incorporation as a regularization mechanism.

Method	Precision	Recall	F1-Score	ROC-AUC	Latency (min)
Hybrid (Proposed)	0.893 ±0.028	0.807 ±0.041	0.847 ±0.031	0.912	4.2±1.1
LSTM	0.874 ±0.035	0.777 ±0.048	0.823 ±0.038	0.896	3.8±1.3
XGBoost	0.831 ±0.042	0.789 ±0.051	0.809 ±0.044	0.883	5.1±1.8
Bayesian-only	0.823 ±0.047	0.762 ±0.053	0.791 ±0.048	0.871	5.7±2.0
Fuzzy-only	0.789 ±0.051	0.737 ±0.058	0.762 ±0.052	0.834	6.3±2.4
Traditional MNF	0.724 ±0.061	0.643 ±0.072	0.681 ±0.064	0.763	180-360

Table IX Test Set Performance Comparison (5-Fold Cv, N=3,240 Samples)

Four observations emerge. False positive reduction: the hybrid achieves a 34% reduction relative to MNF (precision 0.893 vs. 0.724), with 95% CI [27%, 41%] ($p = 0.0037$), translating into reduced field verification costs. Small leak detection: recall improvements over LSTM (0.807 vs. 0.777) indicate enhanced detection of leaks <5 L/min, where Bayesian belief propagation leverages temporal consistency to accumulate evidence that single-snapshot classifiers miss. Detection latency: 4.2 ± 1.1 minutes versus 3–6 hours for MNF, enabling proactive leak management. Data efficiency: although LSTM achieves competitive ROC-AUC (0.896 vs. 0.912), it requires approximately 10× more training data; the hybrid maintains F1 = 0.847 with only 428 labeled hours, demonstrating the value of

Ablation Study

Ablation experiments systematically isolated component contributions. Table X presents results for five configurations.

Configuration	F1-Score	ΔF1 vs. Hybrid	Interpretation
Full Hybrid ($\prime=0.4$)	0.847	—	Baseline
Bayesian-only ($\prime=0$, P(LS) threshold)	0.791	-0.056	Fuzzy adds +5.6% via epistemic modeling
Fuzzy-only ($\prime=1$, HVI threshold)	0.762	-0.085	Bayesian adds +8.5% via probabilistic reasoning
No SNR weighting (SR=1 constant)	0.819	-0.028	Signal quality metrics add +2.8%
Static $\prime=0.5$ (not optimized)	0.831	-0.016	Optimized \prime adds +1.6%

Table X Ablation Study: Component Contributions

Bayesian reasoning contributes the largest individual gain (+8.5% F1 over fuzzy-only), consistent with its capacity to integrate temporal evidence and calibrate posterior probabilities. Fuzzy modeling contributes +5.6% over Bayesian-only, reflecting its role in encoding operator expertise that compensates for limited labeled data. SNR-based confidence weighting adds +2.8%, confirming that explicit sensor reliability modeling improves robustness. The optimized fusion weight ($\prime=0.4$ vs. static 0.5) adds +1.6%, suggesting robustness to exact fusion parameterization a desirable deployment property.

The cumulative decomposition (5.6% + 8.5% + 2.8% + 1.6% = 18.5%) exceeds the observed total improvement over MNF ($\Delta F1 = 0.166$), indicating positive interaction effects: components reinforce each other non-additively, with the fuzzy layer providing structured inputs that improve Bayesian convergence.

Pilot Network Evaluation

The hybrid model was evaluated on the Coleipa pilot network under controlled conditions with clearly distinguishable leak signatures. The 360 labeled samples (305 normal, 55 leak events) were processed through the complete pipeline. Classification performance was assessed via confusion matrix (Table XI) and classification report (Table XII).

	Predicted Normal	Predicted Leak
Actual Normal	305 (TN)	0 (FP)
Actual Leak	0 (FN)	55 (TP)

Table XI Confusion Matrix (Pilot Network, N = 360)

	Precision	Recall	F1 score	support
Normal	1.00	1.00	1.00	305
Leakage	1.00	1.00	1.00	55
Accuracy			1.00	360
Macro avg	1.00	1.00	1.00	360
Weighted avg	1.00	1.00	1.00	360

Table XII Classification Report (Pilot Network)

The model achieved perfect classification (accuracy = 1.00, F1 = 1.00) with zero false positives and false negatives, reflecting the distinct separability of leak profiles under pilot sector conditions. The controlled

valve openings produced flow deviations of 3.2–18.7 L/min against background noise of approximately 2.1 L/min. While these results validate the architectural logic and sequential coupling between fuzzy and Bayesian modules, they should not be extrapolated to larger, more meshed networks with higher background noise. The cross-validated results (F1 = 0.847) provide a more conservative and generalizable estimate.

Traditional methods demonstrated acceptable performance under stable conditions but exhibited limitations under irregular consumption patterns [1], [5]. Acoustic detection was hindered by environmental noise (SNR < 10 dB during peak hours), while hydraulic modeling required extensive calibration not feasible given available infrastructure data. The hybrid model adapted dynamically to variations, leveraging both expert rules and probabilistic updating to maintain reliable detection even with incomplete inputs.

Failure Mode Analysis

The test set produced 7 false negatives and 12 false positives across two dominant failure modes. False negatives (n = 7) clustered in: (1) small leaks (<3.5 L/min) during high-demand periods (07:00–09:00), masked by demand fluctuations ('_demand = 4.2 L/min, SNR < 8 dB) a demand forecasting module (ARIMA or LSTM-based) could increase effective SNR; (2) gradual leak onset (<0.5 L/min per hour), where Bayesian prior adaptation lagged behind evidence accumulation a CUSUM-based drift detection mechanism could address this. False positives (n = 12) occurred during abrupt demand changes (valve operations, hydrant flushing) mimicking leak signatures without SCADA event log entries. Inte-

gration with utility work order management systems could suppress approximately 75% of these false alarms. These failure modes define actionable improvement vectors and do not compromise the overall performance advantage.

Discussion

Architectural Advantages

The architecture provides three distinct advantages over prior detection algorithms. First, unlike black-box models, it propagates uncertainty metrics fuzzy membership degrees and Bayesian credible intervals throughout the supervisory pipeline, enabling operators to receive quantified confidence levels for risk-adjusted response prioritization. A high-risk alert driven primarily by the Bayesian component signals strong data-driven evidence, whereas a fuzzy-dominant alert indicates expert heuristics compensating for weak sensor evidence.

Second, small-sample robustness addresses a deployment barrier in under-resourced utilities. The hybrid leverages domain knowledge and prior beliefs to achieve high performance with a fraction of the data required by purely neural approaches, effectively regularizing inference under labeled data scarcity. Third, interpretability enables operators to trace decisions through the processing chain via transparent explanations correlating hydraulic deviations with historical signatures, fostering trust and supporting regulatory compliance requirements.

Scalability Validation

EPANET 2.2 simulations on synthetic networks (50 to 1,200 nodes) assessed scala-

bility. Table XIII summarizes computational performance.

Network Size (nodes)	Sensors	Processing Time (ms)	Memory (MB)	F1-Score
50	3	180±22	42	0.839
250	8	420±38	98	0.841
600	15	680±51	178	0.836
1,200	28	840±67	312	0.831

Table XIII Scalability Analysis On Simulated Networks (Epanet 2.2)

At 1,200 nodes, processing time remains below 850 ms within the 1-minute SCADA sampling interval. The approximately linear scaling ($R^2 = 0.98$) indicates the bottleneck resides in Bayesian belief propagation ($O(n \cdot k^2)$), not fuzzy inference (constant-time per zone). Memory scales linearly at ~0.26 MB/node, enabling deployment on commodity hardware (312 MB for 1,200 nodes). F1-score degradation from 50 to 1,200 nodes is modest (-1.9%), attributed to increased sensor noise propagation. Hierarchical architectures partitioning into DMAs with independent hybrid instances are recommended for networks exceeding 500 nodes.

Bridging Research and Field Operations

The case study provides empirical evidence that hybrid approaches can outperform conventional techniques, particularly where instrumentation is limited and operational variability is high [6]. The SCADA implementation demonstrated operational viability: dashboards with real-time monitoring, classification outputs, and probabilistic risk indicators enabled effective operator interaction. This adaptability is particularly

relevant for small and medium-sized utilities in developing regions, where infrastructure data are sparse and operational budgets constrain adoption of complex technologies.

Limitations and Future Work

Empirical validation was conducted on a single pilot network with tree-topology and PVC material. While simulations provide scalability evidence, broader field trials across diverse topologies including looped networks and varying materials or climates are essential to confirm generalizability. The current evaluation's dependency on accurate demand forecasting to detect small leaks during high-variability periods suggests that advanced forecasting models trained on multi-year consumption patterns could reduce the false negative rate.

The system currently classifies events as leak or normal without distinguishing burst events from gradual leaks. Augmenting the Bayesian logic with severity nodes or rate-of-change features would better align with operational priority protocols. Additionally, incorporating water quality indicators, energy consumption patterns, and infrastructure metadata could enhance detection capabilities. Future directions include multi-variate sensors for contamination detection, active learning frameworks refining fuzzy rules via operator feedback, federated learning for cross-utility knowledge sharing, and comprehensive usability evaluations with utility staff.

CONCLUSION

This work presented a hybrid Fuzzy–Bayesian supervisory architecture that integrates uncertainty reasoning directly within SCADA data flows. The modular fi-

ve-layer design harmonizes fuzzy linguistic models encoding operator expertise with Bayesian belief networks for probabilistic reasoning, enabling explicit confidence propagation from raw sensor inputs to operator dashboards.

Validation on a pilot network in Coileipa, Santa Bárbara do Pará, Brazil (360 operational hours, SNR = 12.3 dB) achieved F1-score = 0.847 ± 0.031 under 5-fold cross-validation, with a 34% false positive reduction versus MNF analysis (95% CI: [27%, 41%], $p < 0.01$). The hybrid outperformed standalone fuzzy (+8.5% F1), Bayesian (+5.6% F1), LSTM, and XGBoost baselines while requiring 96% fewer training samples. Under controlled conditions, perfect class separation was achieved (F1 = 1.00). Scalability tests via EPANET 2.2 confirmed processing latency below 850 ms on networks up to 1,200 nodes.

Four contributions distinguish this work: (1) architectural fusion of epistemic and aleatoric uncertainty within supervisory control loops; (2) small-sample robustness via fuzzy rules and Bayesian priors; (3) operator-interpretable outputs with decomposed confidence metrics; (4) field validation on a network representative of under-resourced urban systems in developing regions.

Future research will incorporate water quality sensors for contamination detection, federated learning for cross-utility knowledge transfer, and burst severity classification. Open-source implementation and datasets are available at <https://github.com/vitor-zener/deteccaovazamento>.

Acknowledgment

This study was financed in part by the CAPES. The authors also thank the local water utility for providing access to data and operational support during field validation, and the local utility staff of the Coleipa pilot network for their cooperation in data collection and sensor maintenance. Contributions from utility engineers J. Silva, M. Santos, and R. Oliveira for fuzzy rule elicitation via Delphi interviews are gratefully acknowledged.

References

1. N. Ullah, Z. Ahmad, and J.-M. Kim, "Pipeline leakage detection via anomaly detection and machine learning algorithms," *Sensors*, vol. 23, no. 6, Art. no. 3226, 2023, doi: 10.3390/s23063226.
2. E. Farah and I. Shahrour, "Water leak detection: A comprehensive review of technologies, challenges, and future directions," *Water*, vol. 16, no. 20, Art. no. 2975, 2024, doi: 10.3390/w16202975.
3. A. Momeni, M. L. A. da Fonseca, and A. H. Reis, "Hydraulic model-based leakage detection in water distribution networks using pressure monitoring data," *Frontiers in Water*, vol. 3, Art. 648622, 2021, doi: 10.3389/frwa.2021.648622.
4. L. Romero-Ben, D. Alves, J. Blesa, G. Cembrano, V. Puig, and E. Duviella, "Leak detection and localization in water distribution networks: Review and perspective," *Annual Reviews in Control*, vol. 55, pp. 392–419, 2023, doi: 10.1016/j.arcontrol.2023.03.012.
5. I. Daniel, S. Meseguer, J. Blesa, and V. Puig, "A sequential pressure-based algorithm for data-driven leak localization in water distribution networks," *J. Water Resour. Planning Manage.*, vol. 148, no. 11, 2022, Art. 04022059, doi: 10.1061/(ASCE)WR.1943-5452.0001535.
6. A. Soldevila, R. M. Fernandez-Canti, J. Blesa, S. Tornil-Sin, and V. Puig, "Leak localization in water distribution networks using Bayesian classifiers," *J. Process Control*, vol. 55, pp. 1–15, 2017, doi: 10.1016/j.jprocont.2017.03.015.
7. Z. Y. Wu, A. Chew, X. Meng, J. Cai, and J. Pok, "High fidelity digital twin-based anomaly detection and localization for smart water grid operation management," *Sustainable Cities and Society*, vol. 95, 2023, Art. 104446, doi: 10.1016/j.scs.2023.104446.
8. B. Henriques-Silva, A. E. S. Costa, and S. M. Gomes, "Burst warning system for water distribution networks based on Bayesian dynamic linear models," *Water Resources Research*, vol. 59, no. 7, 2023, e2022WR033550, doi: 10.1029/2022WR033550.
9. D. Barros, I. Almeida, A. Zanfei, and G. C. G. Garcia, "An investigation on the effect of leakages on the water quality parameters in distribution networks," *Water*, vol. 15, no. 2, Art. no. 324, 2023, doi: 10.3390/w15020324.
10. M. B. Moubayed and H. Sharif, "Water leak detection survey: Challenges & research opportunities," *IEEE Access*, vol. 9, pp. 1–22, 2021, doi: 10.1109/ACCESS.2021.3112192.
11. A. Plumstead and P. Samson, "SCADA master planning for water and wastewater utilities," *J. Amer. Water Works Assoc.*, vol. 115, no. 8, pp. 12–21, 2023, doi: 10.1002/awwa.2158.
12. L. A. Rossman, *EPANET 2: Users Manual*.

Cincinnati, OH, USA: U.S. Environmental Protection Agency, 2000.

13. K. A. Klise, R. Murray, and T. Haxton, "Water Network Tool for Resilience (WNTR): Users manual and software," Sandia National Laboratories, 2018.
14. L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965, doi: 10.1016/S0019-9958(65)90241-X.
15. T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.
16. I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. Cambridge, MA, USA: Morgan Kaufmann, 2016.
17. R. Puust, Z. Kapelan, D. A. Savic, and T. Koppel, "A review of methods for leakage management in pipe networks," *Urban Water Journal*, vol. 7, no. 1, pp. 25–45, 2010, doi: 10.1080/15730621003610878.
18. M. Sant'Anna, A. M. Mendiondo, and E. E. A. Righetto, "Analysis of the impact of leaks and valve maneuvers on pressure signals using frequency-domain indicators," *RBRH*, vol. 29, 2024, e9, doi: 10.1590/2318-0331.292420240090.
19. A. S. O'Hagan and J. J. Forster, *Kendall's Advanced Theory of Statistics, Vol. 2B: Bayesian Inference*, 2nd ed. London, U.K.: Arnold, 2004.
20. G. V. S. H. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River, NJ, USA: Prentice Hall, 1995.
21. R. A. Cody, B. A. Tolson, and J. Orchard, "Detecting leaks in water distribution pipes using a deep autoencoder and hydroacoustic spectrograms," *J. Infrastructure Preservation and Resilience*, vol. 2, no. 1, pp. 1–13, 2021, doi: 10.1186/s43065-021-00021-6.
22. O. C. Izinyon, U. C. Nwobi-Okoye, and A. O. Igbinoba, "Development of a fault detection and localization model for a water distribution network," *Applied Sciences*, vol. 14, no. 4, Art. no. 1620, 2024, doi: 10.3390/app14041620.
23. H. M. Tornyeviadzi and R. Seidu, "Leakage detection in water distribution networks via 1D CNN deep autoencoder for multivariate SCADA data," *Eng. Appl. Artif. Intell.*, vol. 122, Art. 106062, 2023, doi: 10.1016/j.engappai.2023.106062.
24. J. Choi and S. Im, "Application of CNN models to detect and classify leakages in water pipelines using magnitude spectra of vibration sound," *Applied Sciences*, vol. 13, no. 5, Art. no. 2845, 2023, doi: 10.3390/app13052845.
25. G. Örn Garðarsson, F. Boem, and L. Toni, "Graph-based learning for leak detection and localisation in water distribution networks," *IFAC-PapersOnLine*, vol. 55, no. 6, pp. 661–666, 2022, doi: 10.1016/j.ifacol.2022.07.203.
26. X. Delgado-Galván, R. Pérez-García, J. Izquierdo, and J. Mora-Rodríguez, "An analytic hierarchy process for assessing externalities in water leakage management," *Math. Comput. Model.*, vol. 52, no. 7–8, pp. 1194–1202, 2010, doi: 10.1016/j.mcm.2010.03.014.
27. G. van Lagen, E. Abraham, and P. Mohajerin Esfahani, "A Bayesian approach for active fault isolation with an application to leakage localization in water distribution networks," *IEEE Trans. Control Syst. Technol.*, vol. 31, no. 2, pp. 614–629, 2023, doi: 10.1109/TCST.2022.3205438.

28. A. Soldevila, G. Boracchi, M. Roveri, S. Tornil-Sin, and V. Puig, “Leak detection and localization in water distribution networks by combining expert knowledge and data-driven models,” *Neural Comput. Appl.*, vol. 34, no. 6, pp. 4759–4779, 2022, doi: 10.1007/s00521-021-06666-4.
29. J. M. Kościelny, M. Syfert, and P. Wnuk, “Diagnosing with a hybrid fuzzy–Bayesian inference approach,” *Eng. Appl. Artif. Intell.*, vol. 104, Art. 104375, 2021, doi: 10.1016/j.engappai.2021.104375.
30. D. P. Sousa, R. Du, J. M. B. da Silva Jr., C. C. Cavalcante, and C. Fischione, “Leakage detection in water distribution networks using machine-learning strategies,” *Water Supply*, vol. 23, no. 3, pp. 1115–1126, 2023, doi: 10.2166/ws.2023.054.