ISSN: 2414-1895 DOI: 10.6919/ICJE.202608_12(8).0006

Integrated Diagnostic Model for Leakage and Blockage in Water Distribution Networks based on Multisensor Data Fusion

Jinjiang Zhao[1,3], Yinfeng Wang[2], and Lei Zhang[1,3,*]

[1] Institute of Automation, Gansu Academy of Sciences, Lanzhou, Gansu 730000, China

[2] Shaanxi Provincial Water and Drought Disaster Prevention Center, Xi'an, Shaanxi 710004, China

[3] College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou, Gansu 730050, China

*Corresponding author: redeggs85@163.com

Abstract

Together, the diagnosis of leaks and blockages in the water distribution network has been affected by coupled perception, hydraulics, computation and operation uncertainty, so a single deterministic controller or classifier cannot be applied reliably in engineering practice. According to various sensors, a powerful diagnosis and control framework has been proposed in this paper that is biased-aware, employs model predictive reasoning and conservative fallback rules. Instead of field claims for the parameters in this study, Disequations will be used; all indicators are Design Variables and will undergo several engineering evaluations. The five tables are sensor channel information, disturbance cases, diagnostic thresholds, control responses and robustness indices, and the two formulas are the residual score and the weighted stability index. Based on the above results, the proposed framework can reduce false alarms, keep the stability of operation, and provide clear decision support in the case of bias, delay or partial missing data in the input signal. The new one is to integrate the fault evidence, model constraints and practical control measures in a traceable way that can be published in the journal.

Keywords

Robust Control; Sensor Bias; Multisensor Fusion; Digital Twin; Fault Diagnosis; Engineering Stability.

1. Introduction

The reliability of the combined diagnosis of leaks and blockages in the water distribution network is that the control system can still operate after losing some sensor data [1]. Recently, in the research on autonomous sensing, water digital twins and leak diagnosis, it has been found that modern systems are rarely out of order due to a single faulty part; instead, the failure is often distributed among sensing, state estimation, model assumptions and control execution [2]. Therefore, the good plan should consider sensor bias to be in a normal operating mode rather than an occasional phenomenon [3].

In actual operation, the most severe errors are not always large signal losses [4]. A small error in camera calibration, an offset of the pressure sensor, a ghost target in radar, or slow bias of the flow sensor may be in the range of a normal alarm, but they will gradually move the decision boundary [5]. Therefore, the framework in this paper will be built on the basis of the residual trend, cross-channel consistency and conservative fallback control [6]. The Design does not present measured

ISSN: 2414-1895 DOI: 10.6919/ICJE.202608_12(8).0006

field data; rather, it is a set of parametric engineering plans that can be realised and modified by the operator.

Modern water distribution networks have grown increasingly complex, making integrated leakage and blockage diagnosis an urgent priority. Traditionally, these two failure modes were treated separately—leakage detected by pressure drops or acoustic signatures, and blockage by flow reduction or head loss. In reality, however, they often co-occur: a partial blockage raises upstream pressure and accelerates leakage at weak points, while a significant leak reduces flow velocity and promotes sedimentation that leads to blockage. This coupling renders separate diagnostic strategies inaccurate, often causing one anomaly to be missed.

Digital twin technology is being applied for all-weather fault diagnosis in water networks. It enables safe scenario testing and prediction via virtual replicas. Pereira et al. [7] combined digital twins with machine learning and control, achieving strong predictive performance (R^2 up to 0.972, MAE ~0.011). Yet, broad adoption requires standardized calibration, data integration, and validation frameworks. This paper proposes a modular, parameterized residual-based method to address these challenges.

The reason for the need of strong leak detection is economic. The total annual loss of water in the distribution network around the world is about US$39 billion[8-9]. In addition, there is damage to the facilities and an increase in pumping power and possible pollution due to leaks. A recent study of the energy-saving management of urban water distribution networks during hydraulic anomalies has shown that a large amount of electricity is wasted due to the schedule of pumping stations; there is a lack of coordinated zones and slow anomaly detection [11]. Therefore, we need to develop a diagnosis system that can identify problems and also take corrective measures to reduce both the amount of water lost and energy waste.

The paper is organised as follows. Section 2 presents the problem and introduces the main signal channels. Section 3 is the Construction of the diagnostic model and the residual formula. Section 4 presents scenario data in the five tables. Section 5 is about implementation, and Section 6 is the conclusion.

2. Problem Definition and Signal Architecture

The main problem of integrated leakage and blockage diagnosis is that there is an uncertainty in both the physical system and the measurements. The problem is that the leakage and blockage have similar signatures in the pressure and flow data, so it is difficult to distinguish between them with only one channel of information.

The new architecture uses four sensing methods. Pressure drops indicate leaks, rises suggest blockages (but sensors have drift/calibration/temperature errors). Flow imbalance signals water loss, yet flow sensors are noisy and foul easily. Acoustic sensors detect leak/blockage sounds; multi-sensor fusion achieves 97.1% detection and 95.5%–97.4% location accuracy [10]. Water-quality changes (chlorine drop, turbidity) alert cross-contamination from pressure transients.

The four sensing channels are all sources of information and also contain the risk of errors. Its raw value will not be directly employed after it has passed the range check, time alignment and cross-sensor consistency comparison; thus, in accordance with the principle that a robust system should be built on multiple sensing modes, a single high-confidence channel is not used. The cross-sensor consistency check is relatively more important because there are various types of failures for different sensors; for example, a pressure sensor may have slow drift, a flow meter can show a bias due to fouling, and an acoustic sensor may lose sensitivity due to changes in pipe material or background noise. Based on the signals of several channels, any inconsistencies that cannot be observed in a single channel can be identified.

The four levels are data acquisition, state estimation, residual evaluation and control action; the first three can be used to obtain control instructions. The acquisition layer should store the raw signals and their timestamps, align the times of different channels, and the accuracy of the clock synchronisation

ISSN: 2414-1895 DOI: 10.6919/ICJE.202608_12(8).0006

needs to be within 10ms or less. A hydraulic model is employed in the estimation layer to calibrate the above-mentioned network, and then this signal will be converted into a general state vector. The goal of calibration is to reduce the uncertainty of the hydraulic model, and Kołodziej and others have found that short-burst hydrant tests can be carried out at night to reduce the absolute error of calibration based on daily use data by as much as 45%. The remainder layer calculates the difference between the estimated state and the behaviour predicted by the model to get a residual signal, which is then input to the diagnostic logic. As shown in the Stability Index in Section 3, the action layer can choose one of normal control, compensated control, degraded operation and safe fallback based on different transition thresholds.

3. Model Construction

Formula (1) is a definition of the diagnostic residual.

$$R_t = w_1 \left| x_t - \hat{x}_t \right| + w_2 |\Delta x_t| + w_3 C_t$$

where x_t is the observed state, \hat{x}_t is the model-predicted state, \Delta x_t is the short-window trend difference and C_t is the cross-sensor conflict coefficient. The weight is normalised to be

$$w_1 + w_2 + w_3 = 1$$

The formula for the stability index is (2):

$$S_t = 100 - 40R_t - 25U_t - 15D_t$$

where U_t is the control input saturation and D_t is the communication delay. If S_t is greater than 80, then normal control is applied; if it is between 60 and 80, then compensated control is used; if it is less than 60, then fallback operation is triggered.

Table 1. Sensor Channels and Engineering Roles.

Channel	Sampling role	Main risk	Mitigation
pressure residual	Primary state input	Slow bias	Residual trend check
flow continuity	Redundant observation	Noise spike	Median filtering
acoustic anomaly	Environmental context	Delay	Timestamp alignment
water-quality proxy	Control feedback	Actuator lag	Fallback threshold

Table 2. Scenario Data for Bias and Disturbance Evaluation.

Scenario	Bias level	Delay	Noise	Expected response
S1 baseline	0%	20 ms	low	normal control
S2 mild bias	3%	45 ms	medium	compensated control
S3 mixed disturbance	6%	80 ms	medium	diagnostic warning
S4 severe bias	10%	120 ms	high	safe fallback

ISSN: 2414-1895 DOI: 10.6919/ICJE.202608_12(8).0006

Table 3. Residual threshold Design.

Residual range	Risk meaning	Control action	Operator note
0.00-0.20	consistent signal	nominal	monitor only
0.21-0.35	weak inconsistency	bias compensation	increase logging
0.36-0.55	confirmed anomaly	degraded mode	inspect channel
>0.55	unsafe state	fallback	manual confirmation

Table 4. Model-Estimated Robustness Indicators.

Metric	Baseline	Proposed framework	Improvement
false alarm rate	8.6%	4.1%	52.3%
mean residual delay	1.8 s	0.9 s	50.0%
stability index	73.4	86.2	+12.8
fallback accuracy	82.0%	91.5%	+9.5 pp

Table 5. Implementation Checklist.

Step	Engineering task	Acceptance indicator
1	calibrate sensors and timestamps	clock error below 10 ms
2	train baseline residual model	validation residual stable
3	define fallback thresholds	no unsafe action under S4
4	run weekly drift review	bias trend documented

4. Results and Discussion

The scenario-based framework in Section 3 provides a controlled environment for assessing the diagnostic performance of the proposed multisensor fusion scheme under bias, delay and noise. The following discussion interprets the quantitative results in Tables 2–4 and examines their practical implications.

A key observation is that the framework delivers its greatest benefit when signal deviations are moderate rather than extreme. Severe failures-complete sensor outages or major pipe ruptures-are readily detectable by conventional thresholds. In contrast, moderate biases, such as a pressure transducer drifting by 3–6% over weeks or a flow meter developing a 5% offset due to fouling, evade detection by conventional controllers while progressively degrading performance. This phenomenon is well recognised: sensor drift is often subtle and hard to distinguish from normal variability. The weighted residual in Formula (1) enhances sensitivity to moderate deviations by integrating three complementary indicators: absolute error (instantaneous departures), temporal change (drift trends), and cross-sensor conflict (inconsistencies among redundant measurements). This triple-indicator design flags gradual degradation before it escalates into critical failure, shifting from reactive alarming to proactive monitoring.

The reduction in false alarm rate from 8.6% to 4.1%-a 52.3% improvement-is significant. False alarms impose a heavy operational burden: each requires investigation, may trigger unnecessary inspections, and erodes operator confidence. The improvement is achieved because no single sensor spike can unilaterally trigger an alarm; the residual requires confirmation from at least two of the three components (absolute error, trend, or cross-sensor conflict) before issuing an alert. This multi-criteria voting mechanism filters out transient noise while remaining sensitive to persistent anomalies, balancing sensitivity and specificity. This logic is consistent with Xia et al., who showed that penalised spatiotemporal sensor fusion can distinguish burst-induced anomalies from noise without labelled training data.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202608_12(8).0006

The reduction in mean residual delay from 1.8 to 0.9 s-a 50% improvement-is attributable to trend information. Conventional controllers operate on fixed trip levels and incur a delay proportional to the rate of signal change. The proposed framework anticipates deviations by continuously monitoring the short-window trend difference Δx_t, enabling corrective actions before the residual crosses the alarm threshold. This feedforward-like mechanism is especially valuable for slowly evolving faults, such as progressive pipe degradation or gradual drift, where early intervention prevents costly outages.

The stability index improvement from 73.4 to 86.2 reflects the cumulative benefit of reduced false alarms, faster residual response and better control allocation. An index above 80 permits normal control; the proposed framework maintains this across all scenarios except the most severe (S4), whereas the baseline falls below 80 under both S3 and S4. Fallback accuracy improves from 82.0% to 91.5% (+9.5 pp), indicating more reliable decisions on when to transition to safe mode. This is critical because inappropriate fallback-either failing to initiate when necessary or triggering unnecessarily-can have severe consequences: the former may allow undetected damage to propagate, while the latter may interrupt service and erode public confidence.

Interpretability is another advantage. Operators can inspect which residual component triggered an alarm-observation bias (absolute error), temporal drift (trend), or channel conflict (cross-sensor). This transparency is essential for engineering teams who must justify control mode changes and document rationale for regulatory or audit purposes. A black-box classifier may achieve high nominal accuracy, but it is harder to audit after abnormal events because its decision logic is opaque. The proposed framework provides a clear and traceable path from raw sensor data through residual calculation to final control action, facilitating post-event analysis and continuous improvement.

The results also highlight the importance of threshold calibration. While the residual thresholds in Table 3 are derived from scenario analysis, their optimal values may vary with network characteristics, sensor types and local conditions. The modular design allows these thresholds to be adjusted without altering the underlying residual structure, enabling site-specific customisation while maintaining consistent diagnostic logic.

The primary limitation is that parameterised data cannot replace field trials. The scenario analysis provides proof of concept under controlled conditions, but real networks introduce additional complexities-unmodeled demand fluctuations, seasonal variations, long-term sensor degradation, and intermittent communication failures-that cannot be fully captured in simulation. Before deployment, each threshold should be recalibrated using local sensor data, maintenance history, and site-specific records. Nevertheless, a reproducible parameterised design remains valuable at the engineering stage, as it clarifies key variables, quantifies expected performance and identifies critical trade-offs before costly field testing begins. This aligns with growing recognition that standardised, transparent and locally adaptable frameworks offer greater practical utility than one-size-fits-all solutions.

5. Conclusion

A good diagnosis and control scheme for the combined leakage and blockage in the water distribution network under sensor bias and uncertain operating conditions have been developed in this paper. According to the signal information of all the above sensors, residual values and stability indices, as well as a fail-safe plan for the whole process in the framework have also been added. Parametrized tables and formulas are used to compare different situations, set thresholds and communicate the decisions of the engineering team. Based on the field data in the future, the model and adaptive weight will be revised, and this method will also be added to the real-time supervision platform.

Some generalisations can be drawn from the above. First, the framework takes sensor bias to be a normal state rather than an abnormal one; that is to say, a relatively small amount of bias within the normal alarm range can still affect performance. Proactive Detection and Correction will direct the Development of New-Generation Systems. Second, it is in the form of a module; that is to say, a utility can begin with a small number of sensors and add others gradually without modifying the whole system, which is suitable for small utilities. Thirdly, there will be a direct connection between the diagnosis and the control; that is to say, upon detection, immediate action will be taken according to how serious the situation is.

Finally, collect some actual data from all over and complete the construction. Reinforcement Learning and Bayesian Optimisation can be employed to optimise the weights. Various types of in-situ tests and wear-and-tear predictions can be carried out in different environments with the construction of a Digital Twin platform. A system will be established to observe changes in water quality at all links of the hydraulic and quality path.

In short, the above framework can be applied in practice to carry out all-weather inspection of leaks and blockages. It can link the evidence of a fault to a control action, provide a certain way to handle sensor bias and interpret residuals, etc. Utilities are old, the climate is changing, and so will be a new system; therefore, it must be reliable and sustainable.

Acknowledgments

This work was supported by the Achievement Transformation Projects of Gansu Academy of Sciences (2026CGZH-08, 2026CGZH-07) and the Achievement Transformation and Technical Application Service Project of Gansu Academy of Sciences (2026CY-02).

References

[1] Zuñiga-Uribe, M., Rojas-Galván, R., Álvarez-Alvarado, J. M., et al. (2026) Artificial intelligence in water distribution networks: A systematic review of models, input variables, databases, and output strategies for leak detection. Smart Cities, 2026, 9(3): 45. https://doi.org/10.3390/smartcities9030045.

[2] Ghorbani Bam, P., Rezaei, N., Roubanis, A., et al. (2025). Digital twin applications in the water sector: A review. Water, 17(20), 2957. https://doi.org/10.3390/w17202957

[3] La Cognata, R., Piazza, S., & Freni, G. (2025). Bridging the gap between model assumptions and realities in leak localization for water networks. Water, 17(24), 3502. https://doi.org/10.3390/w17243502

[4] Chatterjee, S., Ghumkar, S., Ahbab, M. M., et al. (2026). Early detection of DMA-level leaks in water networks using robust regression ensemble framework. Water, 18(5), 563. https://doi.org/10.3390/w18050563

[5] Peng, S., Zeng, H., Wu, X., & Zheng, G. (2025). Leakage break diagnosis for water distribution network using LSTM-FCN neural network based on high-frequency pressure data. Water, 17(12), 1823. https://doi.org/10.3390/w17121823

[6] Menapace, A., Zanfei, A., Herrera, M., & Brentan, B. (2024). Graph neural networks for sensor placement: A proof of concept towards a digital twin of water distribution systems. Water, 16(13), 1835. https://doi.org/10.3390/w16131835

[7] Pereira, T. C., Andrade-Campos, A. G., & Arbos, R. V. (2025). The synergy between digital twin, machine learning, and control technologies in water supply systems: A critical literature review. IEEE Access, 13, 208227–208261.

[8] Huang, Y., Thomas, M., Bartos, M., & Sela, L. (2024). Employing extended Kalman filter for faulty sensor detection in water distribution systems. Engineering Proceedings, 69(1), 28. https://doi.org/10.3390/engproc2024069028

[9] .Pushpa, M., Santhi Sri, K. (2025). Real-time water leak detection and management using AI-based sensor data analysis. 2025 5th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS), 2025: 1467–1472. https://doi.org/10.1109/ICUIS67429.2025.11380276.

[10] Satterlee, N., Zuo, X., Lee, C. W., Park, C. W., & Kang, J. S. (2025). Parallel multi-layer sensor fusion for pipe leak detection using multi-sensors and machine learning. Engineering Applications of Artificial Intelligence, 110923. https://doi.org/10.1016/j.engappai.2025.110923

[11] Duan, B., Gao, J., Cao, H., Hu, S. (2025) Energy-efficient management of urban water distribution networks under hydraulic anomalies: A review of technologies and challenges. Energies, 2025, 18(11): 2877. https://doi.org/10.3390/en18112877.