International Journal of Applied Smart Interdisciplinary Technologies (IJASIT)

Home Page: https://ijasit.org **ISSN: 3139-759X**

https://doi.org/10.68104/ijasit.v1.i4.58

Machine Learning-Based Non-Revenue Water Management in Smart Cities: A Systematic Review of Leakage Detection, Prediction and Decision Support

Varsha Katia[1,*], Namira Saiyed[2]

[1 2] Department of Civil Engineering, Faculty of Engineering and Technology, Parul University.
varsha.katia44736@paruluniversity.ac.in, namira.saiyad30770@paruluniversity.ac.

*** Corresponding author. E-mail:**
varsha.katia44736@paruluniversity.ac.in

Article History:
Received Date: 30 September 2026
Revised Date: 14 October 2026
Accepted Date: 27 October 2026
Published Date: 05 November 2026

Keywords:

Water distribution networks; machine learning; smart water; IoT; pressure sensors; acoustic sensing; leak detection; leak localization; non-revenue water; decision support; and predictive maintenance.

ABSTRACT

Water entering a distribution system may be lost physically through leaks, lost economically through metering and billing issues, or used without being invoiced, making Non-Revenue Water (NRW) a recurring concern for metropolitan water utilities. In order to identify anomalous network behavior, locate leaks, and assist with operational decisions, machine learning (ML), the Internet of Things (IoT), advanced metering infrastructure (AMI), hydraulic modeling, and cloud/edge computing are rapidly being integrated. With a focus on leakage detection, localization, prediction, and decision support, this paper summarizes the development of data-driven NRW management. Data sources, pre-processing and feature engineering, algorithm families, validation procedures, and operational deployment are the five dimensions around which the evidence is arranged in the review, which follows an organized systematic-review approach based on PRISMA 2020. In addition to hydraulic-model residuals and network architecture, pressure, flow, acoustic, and AMI data are analyzed. While CNN, LSTM, and graph-based models are increasingly employed for spatial, temporal, and topological patterns, traditional machine learning models like support vector machines, random forests, and gradient-boosting techniques are still helpful when datasets are moderate and manufactured features are instructive. Moving from high-performing laboratory or simulated trials to resilient utility-scale deployment with sensor sparsity, noise, demand unpredictability, network changes, and limited labeled leak occurrences is another important research problem, according to recent studies. Thus, a closed-loop NRW management system that connects sensing, data quality control, machine learning inference, hydraulic validation, field verification, repair priority, and model retraining is proposed in the review. Standardized datasets, leakage-specific assessment methods, cross-network validation, explainable and physics-informed machine learning, cybersecurity, model-drift management, and economic decision support are important research needs. The study comes to the conclusion that machine learning (ML) should not be used as a stand-alone substitute for hydraulic engineering and field practice, but rather as a decision-support layer inside a larger water-loss-control program.

1. Introduction

Urban water distribution networks (WDNs) are intricate socio-technical systems where water must be transported consistently despite fluctuating demand via aging pipes, valves, pumps, reservoirs, and service connections. Non-Revenue Water (NRW), which is typically interpreted within the international water-balance framework as the combination of unbilled allowed use and water losses, is a key indicator of operational inefficiencies. Water losses are typically separated into two categories: real losses, which are mostly caused by leaks from mains, service connections, and storage facilities, and apparent losses, which are linked to unlawful consumption and errors in metering and accounting [1,2].

Water auditing, district metering, minimum nite flow analysis, pressure management, active leakage control, meter management, and field inspection are all included in traditional NRW management. Because leakage is a physical process controlled by hydraulic pressure, pipe condition, material qualities, and network operation, these methods are still essential. However, SCADA, pressure loggers, flow meters, acoustic sensors, smart meters, and Internet of Things devices are increasingly being used by urban utilities to produce high-frequency readings. The generated data volumes present a chance to add automated pattern recognition and predictive analytics to traditional engineering techniques [6,8,14].

Artificial neural networks and similar statistical techniques were shown to be capable of classifying pressure and flow signals in early data-driven leakage studies [4,5]. While more recent research has shifted toward CNN, LSTM, graph neural networks, and hybrid physics-data techniques, later work included wireless sensor networks, support vector machines, random forests, and ensemble learning [7,9,10,13,15–20]. As a result, the literature shows a shift away from rule-based threshold detection and toward systems that are more data-centric and topology-aware.

Despite these advancements, the dataset and validation conditions must be taken into account when interpreting stated model performance. When deployed on a real network with unknown needs, sensor drift, missing data, undocumented pipe alterations, and several simultaneous leaks, a model trained on synthetic hydraulic scenarios may exhibit drastically different behavior. This gap between simulation and reality, as well as the absence of uniform datasets and evaluation procedures, have been highlighted in recent reviews [11,19,20]. Therefore, the key question is not just which machine learning algorithm yields the maximum accuracy, but also under what network, data, sensing, and validation settings an algorithm can enable trustworthy NRW decisions.

Therefore, this review's goals are to: (i) categorize machine learning applications for NRW management; (ii) compare pressure, flow, acoustic, and AMI data streams; (iii) analyze the advantages and disadvantages of the main ML and deep-learning families; (iv) pinpoint validation and deployment obstacles; (v) link leakage prediction and localization with utility decision support; and (vi) suggest a workable closed-loop framework for smart-city NRW management.

2. Review Methodology

PRISMA 2020 reporting guidelines served as the foundation for the evaluation process [16]. The workflow below outlines the repeatable procedure that must be followed and recorded prior to submission because the current article was created as a research-paper draft rather than as a finished bibliographic database export. The publisher and indexing pages for the important works referenced in the manuscript were cross-checked with the current synthesis.

2.1 Search strategy

The following search terms are suggested: ("non-revenue water" OR "water loss" OR leakage OR "pipe burst") AND ("water distribution network" OR "water distribution system" OR "smart water") AND ("machine learning" OR "deep learning" OR "artificial intelligence" OR "neural network" OR "graph neural network"). Scopus, Web of Science, ScienceDirect, IEEE Xplore, SpringerLink, Taylor & Francis, and Google Scholar are recommended databases. For the technology-focused synthesis, the major publication window may be set to 2017–September 2026; however, the conceptual underpinning is established by retaining seminal works on water loss and leakage management conducted before to 2017.

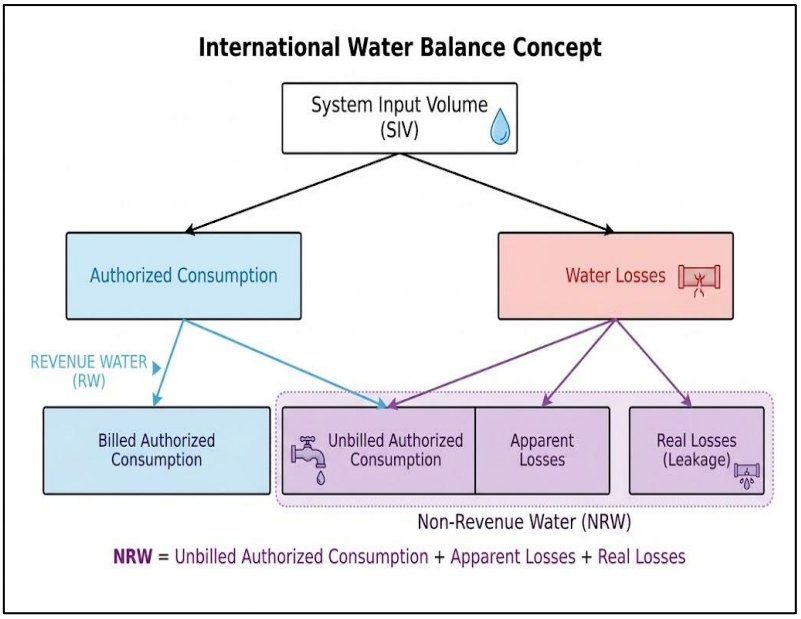


Fig. 1. Conceptual international water-balance structure showing the components that contribute to NRW. The ML focus of this review is mainly on real-loss identification and the data-driven support of apparent-loss control

Table 1
Eligibility criteria

Criterion	Include	Exclude
Topic	Urban potable-water WDNs; NRW; leakage; ML/DL/AI; smart sensing	Oil/gas pipeline studies unless used only as methodological background
Publication	Peer-reviewed journal/conference papers and authoritative technical guidance	Blogs, promotional material, duplicate records
Data	Pressure, flow, acoustic, AMI, SCADA, hydraulic simulation or hybrid data	Studies with no relevance to network water-loss management
Outcome	Detection, localization, prediction, risk assessment, prioritization or decision support	Pure water-quality studies with no NRW/leakage component
Evidence	Studies reporting methodology, data source and validation information	Papers lacking enough methodological detail for interpretation

2.2 Synthesis framework

Five interconnected dimensions are used to analyze each study: (1) sensing and data source; (2) data preprocessing and feature engineering; (3) model family; (4) validation design and performance measures; and (5) operational usefulness. This prioritizes transferability, robustness, and utility value over ranking systems based only on stated accuracy.

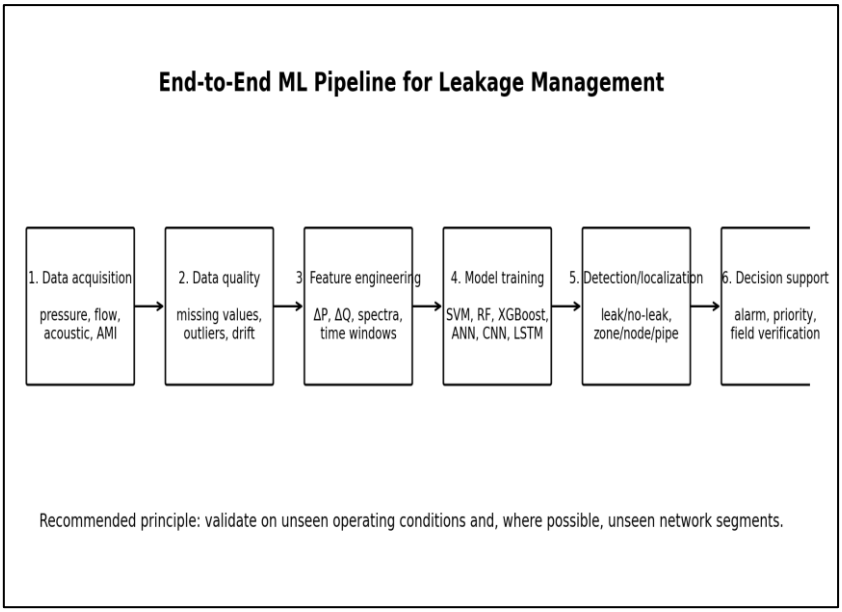


Fig. 2. Recommended evidence-to-deployment pipeline for an ML-based NRW review and implementation framework.

3. NRW and Leakage Management in Smart Cities

3.1 Components of NRW

A consistent foundation for utility auditing is provided by the IWA water-balance concept, which distinguishes between allowed usage and water losses [1]. Unauthorized usage, inaccurate client meters, and mistakes in data handling are examples of apparent losses; leakage from transmission and distribution mains, service connections, and storage are examples of true losses. Because different loss methods require different types of data, the distinction is important for machine learning. Meter and billing data are more suited for apparent-loss analytics, whereas hydraulic pressure and flow data are naturally suited for real-loss and burst detection.

3.2 Why leakage is difficult to manage

Pressure, pipe age, material, diameter, soil conditions, joint conditions, traffic, and construction activity all have an impact on leakage. While sudden bursts can result in significant pressure and flow changes, little background leaks may continue without producing a clear hydraulic signature. Demand variations, especially in highly variable or intermittent systems, can resemble leak characteristics. Therefore, a useful machine learning system needs to be able to discriminate between legitimate operational changes and leakage.

3.3 District Metered Areas and minimum night flow

District Metered Areas (DMAs) continue to be a useful link between smart analytics and traditional water-loss management. Using real-time hydraulic measurements to produce leak alerts, Farah and Shahrour presented a smart-water strategy that combines water balancing with automated minimum nite flow [6]. These methods are useful because they offer an open standard by which more intricate machine learning models may be evaluated.

Table 2
Data Sources and Sensing Modalities

Data type	Typical variables	ML use	Advantages	Main limitations
Pressure	Node pressure, ΔP, pressure residual	Detection, localization, hydraulic-state inference	High temporal value; relatively easy to integrate	Sensor sparsity; pressure-demand confounding
Flow	DMA inflow/outflow,	Water balance, burst detection, forecasting	Direct mass-balance interpretation	Meter accuracy; changing demand

	pipe flow, night flow			
Acoustic	Leak noise, spectra, MFCC, signal energy	Pinpointing and classification	Sensitive to leak signatures	Noise, pipe material, sensor coupling
AMI / smart meters	Customer demand, pressure, temporal consumption	Small-leak detection, demand anomaly	Large spatial coverage	Privacy, data volume, communication reliability
SCADA / operational	Pump status, valve status, tank level	Context and anomaly filtering	Links leak signals with operation	Legacy integration and data quality
Hydraulic simulation	EPANET/WNTR pressure/flow scenarios	Training and scenario generation	Can generate labelled leak cases	Model uncertainty; simulation-to-reality gap

In smart water networks, high-frequency pressure and acoustic detection have become particularly crucial. Applications in online hydraulic modeling, leak detection/localization, and proactive asset management were discovered in a systematic assessment of actual operational networks [14]. However, site-specific implementation and a lack of standardization were noted as significant obstacles.

4. Machine Learning Approaches for NRW Management

4.1 Classical supervised learning

If the labeled datasets are small and feature engineering can reveal hydraulic or acoustic characteristics, methods such as Support Vector Machines, decision trees, random forests, gradient boosting and other ensemble techniques may be of benefit. The effectiveness of these techniques lies in their feature-set development being able to establish the patterns present in the input data. For the purpose of detecting water pipeline leaks, Liu et al. integrated wireless sensing, signal processing, and SVM classification [7]. Using pressure and flow data from EPANET, Mashhadi et al. tested a number of machine learning techniques for leak identification. They found that supervised techniques performed well while unsupervised techniques were more impacted by overlapping clusters [9].

4.2 Artificial neural networks

Nonlinear relationships between measured variables and leak states can be approximated by ANNs. For burst detection, Mounce and Machell employed neural networks in conjunction with hydraulic flow and pressure sensors [4]. For leak identification and localization, more recent lab work has integrated ANN classifiers with designed signal features from accelerometers and dynamic pressure sensors [16]. Nonlinear mapping is the main benefit; sensitivity to training data representativeness is the main drawback.

4.3 Ensemble learning

By merging several weak or somewhat strong learners, ensemble learning can increase robustness. Using several signal characteristics and gradient-boosting models, Ravichandran and colleagues created an ensemble-based acoustic leak classifier that reported less false positives than individual comparison models [10]. Because they may capture nonlinear interactions and offer valuable feature-importance information, tree ensembles are also appealing for tabular sensor data.

4.4 Deep learning: CNN and LSTM

While LSTMs are more suitable for sequential time-series data, CNNs are better suited for spectrograms, pressure-map pictures, and structured spatial representations. DNN, CNN, and SVM models with cloud information management and field operator support have been employed in AI-based acoustic systems [12]. Additionally, a 3D CNN has been created to exploit pressure data from AMI-related events that are spread both spatially and temporally [18]. Under uncertain demand and sensor settings, LSTM-based localization studies classify leak locations using pressure and flow data sequences [17].

4.5 Graph neural networks and topology-aware models

Naturally, a WDN is depicted as a graph with edges representing pipes and nodes representing junctions, tanks, or other hydraulic components. Therefore, instead of treating sensor readings as separate features, GNNs can directly include network topology. Graph-based learning and hybrid hydraulic/ML frameworks, such as

digital-twin implementations that include network mapping, hydraulic calibration, and graph-based leakage identification, are the focus of recent research [16].

4.6 Hybrid and physics-informed ML

When operating conditions diverge from training conditions, models that are solely data-driven may fail. Hydraulic residuals, physical limitations, or digital twin outputs are combined with machine learning in hybrid techniques. This can lessen the quantity of labeled field data needed and increase interpretability. Nevertheless, the hydraulic model, sensor calibration, and network records all introduce uncertainty into hybrid systems. Therefore, the practical goal is to use machine learning as an inference and prioritizing layer rather than to replace hydraulic modeling.

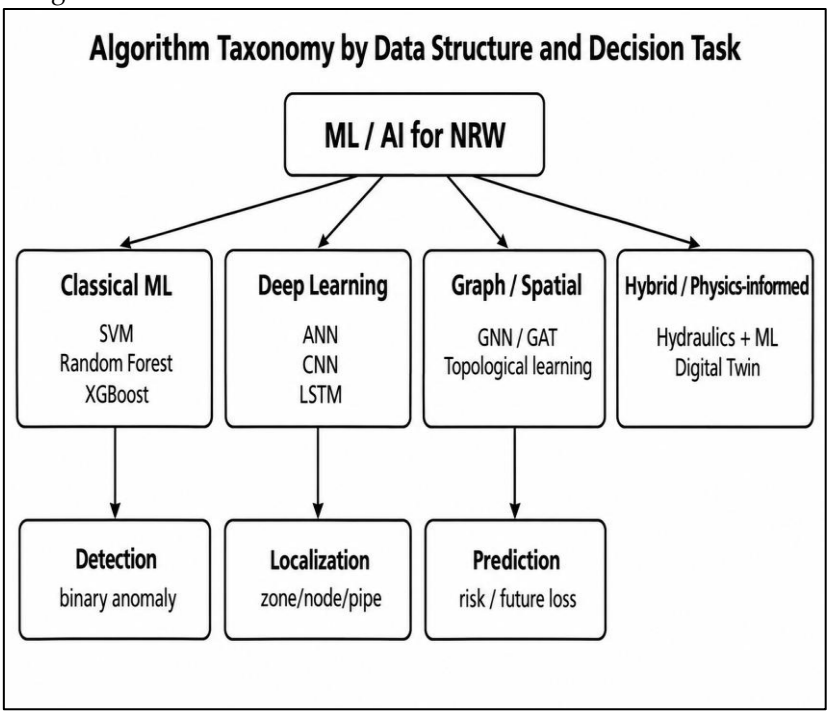


Fig. 3. Taxonomy of ML/AI methods according to data structure and NRW decision task.

5. Leakage Detection, Localization and Prediction

5.1 Detection

Binary classification or anomaly detection—leak vs no-leak—is typically used to develop leak detection. Indicators include changes in minimum nite flow, acoustic signals, and pressure and flow residuals. Binary detection can yield excellent accuracy under controlled conditions, according to recent benchmark research, but this does not ensure dependable field deployment because accuracy is heavily influenced by noise, sensor coverage, and leak size [20].

5.2 Localization

Because the model must determine the location of a leak, frequently from scant measurements, localization is more challenging. Zone classification, node classification, pipe ranking, or continuous coordinate estimation are some possible ways to formulate the job. Comparisons of classifiers have demonstrated that localization is significantly impacted by uncertainty in demand, leak magnitude, and sensor noise [9]. Because they directly depict spatial relationships, graph learning and topology-aware models propose a potential path.

5.3 Prediction and risk assessment

Forecasting leakage risk, forecasting future bursts, evaluating leak magnitude, and identifying assets with elevated failure probability are some related tasks that fall under the category of prediction. Compared to short-term simulated leak scenarios, these jobs demand longitudinal data and asset-condition information, which are less prevalent. Thus, real-time anomaly detection and longer-term asset-risk modeling should be integrated in an advanced smart-city NRW system.

5.4 Multiple and incipient leaks

Small incipient leaks and several simultaneous leaks are still challenging because demand volatility may obscure or overlap their signals. Sensor sparsity, uncertainty and robustness to many leaks are now seen as key evaluation factors in recent research [19, 20]. Utility networks do not work like lab test setups. That is why this shift, in focus is so important.

Table 3

Comparative Synthesis of Representative Studies

Study	Data	Method	Task	Context	Key implication
Mounce & Machell (2006) [4]	Pressure + flow	ANN	Burst detection	Experimental/utility data	Neural classification can learn hydraulic event patterns.
Romano et al. (2014) [5]	Pressure + flow	Wavelets + ANN + SPC + Bayesian inference	Burst/event detection	DMA	Combining AI and statistical control supports near-real-time alarms.
Farah & Shahrour (2017) [6]	Flow + water balance	MNF + thresholds	Leak detection	Campus-scale smart network	Transparent baseline methods remain valuable for ML validation.
Liu et al. (2019) [7]	Sensor signals	SVM + signal processing	Leak detection	WSN + experiments	Feature engineering and low-power sensing are important.
Mashhadi et al. (2021) [9]	Pressure + flow	Multiple ML classifiers	Localization	EPANET + campus	Supervised models can outperform unsupervised clustering for localization.
Ravichandran et al. (2021) [10]	Acoustic	Ensemble gradient boosting	Detection	Field acoustic data	Ensembles can reduce false positives.
Abdelmageed et al. (2022) [11]	Mixed literature	AI review	Detection/localization/prediction/sizing	Review	AI research needs clearer criteria and reproducible comparisons.
Fu et al. (2022) [13]	Urban water data	DL review	Multiple tasks	Review	Data privacy, explainability and digital twins remain major issues.
Rousso et al. (2023) [14]	Pressure + acoustic	SLR of real SWNs	Leak/asset management	Full-scale networks	Real-world evidence is site-specific and insufficiently standardized.
Jun et al. (2024) [18]	AMI pressure	3D CNN	Detection/localization	Real WDN + hydraulic scenarios	Spatiotemporal pressure

					representations are promising but model uncertainty matters.
Farah & Shahrour (2026) [19]	ML/DL literature	Systematic review	Benchmarking practices	Review	Standardized datasets and evaluation protocols remain limited.
Ranganathan et al. (2026) [20]	Hydraulic simulations	ML/DL/GNN/Transformer	Detection/localization	Benchmark WDN	Localization remains more difficult than binary detection; robustness testing is essential.

6. From Prediction to Decision Support

The simultaneous occurrence of multiple leaks and minor leakage points presents a challenge since both types may mask their signals because of the variations in demand. The area of recent studies is shifting its focus towards robustness to multiple leaks, availability of sensors, and operational uncertainty as essential evaluation criteria rather than auxiliary considerations.8.1 Suggested decision logic

1. Screen: find unusual flow, pressure, or sound patterns.
2. Verify by comparing the ML output with the operating condition, nearby sensors, and hydraulic residuals.
3. Localize: rank candidates for DMA, zone, node, or pipe.
4. Set priorities by combining accessibility, repair cost, criticality, probability, and predicted water loss.
5. Verify: update the event label and confirm the field condition.
6. Learn: track model drift and add validated events to the training database.

7. Evaluation Metrics and Validation Practices

For NRW applications, accuracy by itself is insufficient, especially when leakage incidents are uncommon in comparison to typical operating circumstances. If the non-leak class predominates, a model may achieve high accuracy despite overlooking significant leaks. Precision, recall, F1-score, specificity, false-positive and false-negative rates, area under the ROC curve, localization distance/error, and detection time are among the suggested measures. Studies focused on utilities should also include information on the amount of water saved, the length of time that leaks were prevented, the number of field interventions, the cost of the sensor, and the cost of repairs.

Training and test circumstances should be strictly separated during validation. Splitting strongly linked time-series samples at random can result in optimistic estimations and information leakage across sets. Temporal hold-out validation, unseen leak locations, unseen operational conditions, cross-DMA testing, and, when feasible, cross-network validation are examples of more revealing procedures. Before performance results from various research can be effectively compared, consistent datasets and common evaluation methodologies are required, according to current literature [19,20].

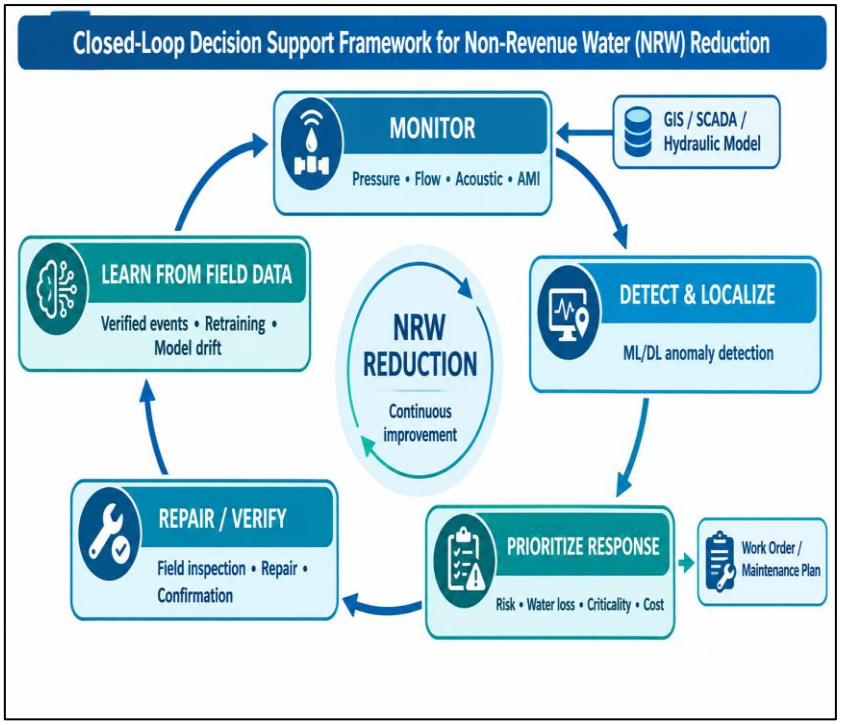


Fig. 4. Proposed closed-loop decision-support process linking monitoring, ML inference, field response and model learning.

Table 4
Major Challenges and Research Gaps

Research gap	Implication / future direction
Data scarcity and label quality	Verified field leaks are comparatively rare, while synthetic data are abundant. Future work should develop shared field datasets, event-label standards and uncertainty-aware learning rather than relying exclusively on simulated scenarios.
Simulation-to-reality transfer	Hydraulic simulators are useful for generating labelled cases, but simplifications in topology, demand, roughness and leak behaviour can cause domain mismatch. Domain adaptation, transfer learning and physics-informed models deserve greater attention.
Sensor sparsity and placement	A utility cannot place high-frequency sensors at every node. Sensor placement should therefore be optimized jointly with model performance, network criticality and monitoring cost.
Noise, missing data and drift	Field sensors experience calibration errors, communication failures and drift. Robust preprocessing and uncertainty-aware models should be evaluated explicitly rather than assuming clean data.
Explainability and engineering trust	Operators must be able to understand why an alarm was triggered and have insight into what evidence supports the suggested location. Explainable AI, confidence intervals, and hydraulic residual visualization can all aid in this process.
Multiple and incipient leaks	Overlapping signatures remain difficult. Future datasets should deliberately include concurrent, small and gradually developing leaks.
Cybersecurity and privacy	Smart meters and connected sensors provide additional attack surfaces for would-be infiltrators and may give away information on customer consumption patterns. Secure data architecture and privacy preserving analytics are therefore pillars of any NRW management strategy.

Economic decision support	The best model is not necessarily the most accurate model. Utilities need to make cost aware decisions considering the value of water saved, the cost of repair and criticality, false alarms, and sensor investment.
Model maintenance	Network topology and operating practices change. Models should include drift monitoring, retraining policies and version control.
Standardization	A common reporting template should specify network characteristics, sensor density, leak sizes, uncertainty, train-test split, metrics, false alarms and field validation.

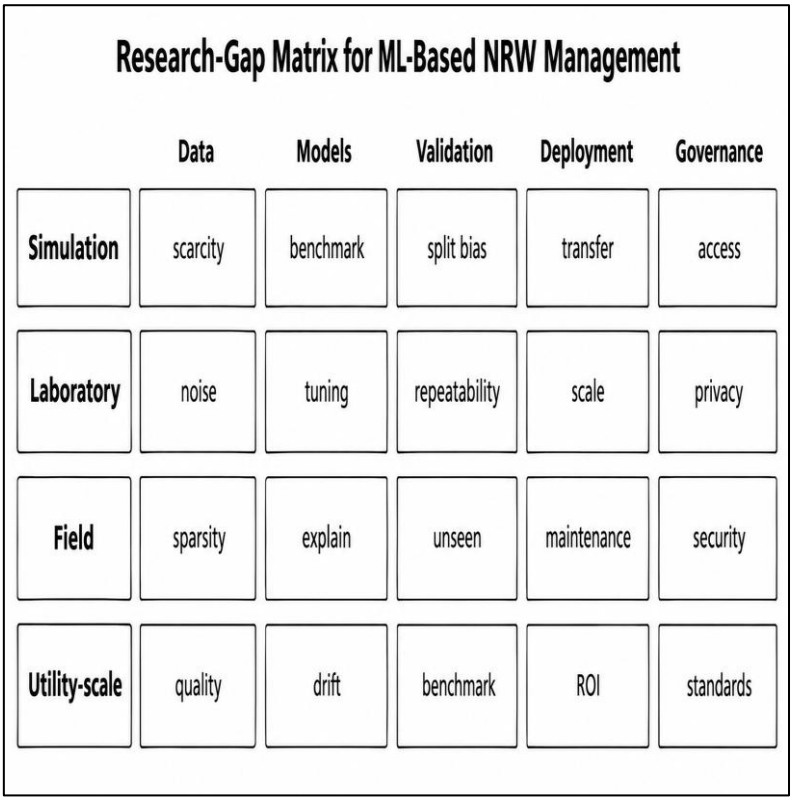


Fig. 5. Conceptual research-gap matrix across data, modelling, validation, deployment and governance dimensions.

8. Proposed Framework for Smart-City NRW Management

A viable design should combine ML and traditional NRW management, instead of considering them as separate strategies, as the literature suggests. The proposed framework has seven layers:

Layer 1 - Asset and hydraulic baseline: retain elements such as GIS, pipe attributes, pressure zones, DMA borders, hydraulic modelling and asset importance.

Layer 2 - Sensing: implement the installation of AMI, flow, pressure and acoustic sensors based on the significance level and danger potential of location.

Layer 3- Data processing: synchronizing time stamps, spotting outliers, estimating missing values, uncertainty measurement, and keeping a record of source.

Layer 4-ML prediction: apply proper techniques for locating, forecasting and determining class and anomalies.

Layer 5: Hydraulic and physical validation: confirm ML notifications according to the functional condition, mass balance, pressure residuals, and nearby readings.

Layer 6: Decision support: classify incidents by the likelihood, seriousness, effect of water leak, importance, and cost of response.

Layer 7: Learning and governance: record validated information, track changes, retrain models, safeguard

cybersecurity, and evaluate performance.

This taxonomy suggests a maturity advancement: level 1 (manual water audits), level 2 (DMA and automatic threshold monitoring), level 3 (supervised machine learning for abnormality detection), level 4 (topology-aware and hybrid models), and level 5 (closed-loop digital decision support with ongoing learning). Each level should show quantified water saving and operational value; it is not compulsory for the utilities to aim for level 5. A maturity progression is implied by this taxonomy: level 1 (manual water audits), level 2 (DMA and automated threshold monitoring), level 3 (supervised machine learning for anomaly detection), level 4 (topology-aware and hybrid models), and level 5 (closed-loop digital decision support with continuous learning). Each stage should exhibit quantified water-saving and operational value; utilities are not obliged to target level 5.

9. Sustainability and Smart-City Implications

The advantages of reducing NRW are the minimization of the amount of water that goes wasted before it is consumed. Yet, the process of digitalization itself can generate costs and consume energy through the use of sensors, connectivity, cloud computing, and maintenance. Thus, sustainability should be seen as a system-level outcome. There exist various useful indicators, such as saved water, the amount of energy supplied per cubic meter, time to react to leaks, sensor lifespan, prevented pumping or treatment, and the return on investment. Interoperability remains important from the point of view of smart cities. Information has to be transmitted using controlled interfaces connecting GIS, SCADA, AMI, work-order systems, hydraulic models, and ML platforms. Digital twins can provide a coordinating layer, but they will be beneficial only if the assets are properly calibrated and constantly updated with real-world data.

10. Future Research Agenda

- Openly formulate, standardized, multi-modal datasets with identified leak tags with pressure, flow rate, sound and AMI data.
- Make time hold-out validation and cross-network validation standard practice.
- Develop the graph neural network and transformer models, which included topology awareness and interpretable engineering features.
- Combine hydraulic equations, physical constraints, and machine learning with hybrid or physics-informed digital twin systems.
- Examine sensor placement as a joint optimization problem that takes network criticality, cost, and information gain into account.
- Create alerts that are aware of uncertainty and report confidence and the anticipated cost of false alarms.
- When raw consumer data cannot be centralized, look into federated or privacy-preserving learning among utilities.
- Integrate ML predictions with maintenance planning and cost optimization rather than stopping at leak classification.

11. Conclusion

Research on Machine Learning is beginning to have an impact on smart water and NRW management, particularly in the identification of anomalous flow, pressure, and sound patterns and the prioritization of field research. Studies show a clear shift from statistical and threshold approaches to supervised machine learning, ensemble learning, and deep learning, as well as topology-aware models. But algorithm creation is no longer the main problem. Reliable data, representative labels, and consistent validation and integration are all part of it. Though acoustic sensors offer more localized evidence, pressure and flow measurements are still vital because there is a direct relationship to hydraulic behavior. AMI enables realistic demand and leakage analyses while allowing the observational coverage to be extended. There are capabilities for spatial, temporal, and topological patterns through CNNs, LSTMs, GNNs, and hybrid models; however, one should be aware of how performance was reported in relation to design of datasets and operating conditions. Hence, the optimal closed-loop scenario for smart cities is when machine learning compliments field engineering, hydraulic modeling, and water-balance auditing practices. The proposed closed-loop scenario involves the connection between

sensing practices, data quality checks, machine learning inference, hydraulic verification, as well as decision support practices, repair, and ongoing learning processes. Future work should investigate frameworks that are not only explainable, and cost-efective, and informed by physical principles, but also tested in the field and capable of addressing complexities such as simultaneous leakage, sparse sensing, uncertainty, and changes in topography.

References

[1]. Lambert, A. O., Brown, T. G., Takizawa, M., & Weimer, D. (1999). A review of performance indicators for real losses from water supply systems. Aqua—Journal of Water Supply Research and Technology, 48(6), 227–237. https://doi.org/10.2166/aqua.1999.0025

[2]. Farley, M., & Liemberger, R. (2005). Developing a non-revenue water reduction strategy: Planning and implementing the strategy. Water Supply, 5(1), 41–50. https://doi.org/10.2166/ws.2005.0006

[3]. Puust, R., Kapelan, Z., Savic, D. A., & Koppel, T. (2010). A review of methods for leakage management in pipe networks. Urban Water Journal, 7(1), 25–45. https://doi.org/10.1080/15730621003610878

[4]. Mounce, S. R., & Machell, J. (2006/2007). Burst detection using hydraulic data from water distribution systems with artificial neural networks. Urban Water Journal, 3(1), 21–31. https://doi.org/10.1080/15730620600578538

[5]. Romano, M., Kapelan, Z., & Savic, D. A. (2014). Automated detection of pipe bursts and other events in water distribution systems. Journal of Water Resources Planning and Management, 140(4), 457–467. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000339

[6]. Farah, E., & Shahrour, I. (2017). Leakage detection using smart water system: Combination of water balance and automated minimum night flow. Water Resources Management, 31, 4821–4833. https://doi.org/10.1007/s11269-017-1780-9

[7]. Liu, Y., Ma, X., Li, Y., Tie, Y., Zhang, Y., & Gao, J. (2019). Water pipeline leakage detection based on machine learning and wireless sensor networks. Sensors, 19(23), 5086. https://doi.org/10.3390/s19235086

[8]. Singh, M., & Ahmed, S. (2021). IoT based smart water management systems: A systematic review. Materials Today: Proceedings, 46, 5211–5218. https://doi.org/10.1016/j.matpr.2020.08.588

[9]. Mashhadi, N., Shahrour, I., Attoue, N., El Khattabi, J., & Aljer, A. (2021). Use of machine learning for leak detection and localization in water distribution systems. Smart Cities, 4(4), 1293–1315. https://doi.org/10.3390/smartcities4040069

[10]. Ravichandran, T., Gavahi, K., Ponnambalam, K., & Mousavi, S. J. (2021). Ensemble-based machine learning approach for improved leak detection in water mains. Journal of Hydroinformatics, 23(2), 307–323. https://doi.org/10.2166/hydro.2021.093

[11]. Abdelmageed, S., Tariq, S., Boadu, V., & Zayed, T. (2022). Criteria-based critical review of artificial intelligence applications in water-leak management. Environmental Reviews, 30(2), 280–297. https://doi.org/10.1139/er-2021-0046

[12]. Vanijjirattikhan, R., Khomsay, S., Kitbutrawat, N., Khomsay, K., et al. (2022). AI-based acoustic leak detection in water distribution systems. Results in Engineering, 15, 100557. https://doi.org/10.1016/j.rineng.2022.100557

[13]. Fu, G., Jin, Y., Sun, S., Yuan, Z., & Butler, D. (2022). The role of deep learning in urban water management: A critical review. Water Research, 223, 118973. https://doi.org/10.1016/j.watres.2022.118973

[14]. Shen, Y., & Cheng, W. (2022). A tree-based machine learning method for pipeline leakage detection. Water, 14(18), 2833. https://doi.org/10.3390/w14182833

[15]. Rousso, B. Z., Lambert, M., & Gong, J. (2023). Smart water networks: A systematic review of applications using high-frequency pressure and acoustic sensors in real water distribution systems. Journal of Cleaner Production, 410, 137193. https://doi.org/10.1016/j.jclepro.2023.137193

[16]. Mahdi, N., Hikmet, A., Abulqasim, S., Basem, A., et al. (2024). Leak detection and localization in water distribution systems using advanced feature analysis and an artificial neural network. Desalination and Water Treatment, 320, 100685. https://doi.org/10.1016/j.dwt.2024.100685

[17]. Gómez-Coronel, L., Santos-Ruiz, I., Blesa, J., Puig, V., & Lopez-Estrada, F. R. (2024). Leak localization in an urban water distribution network using a LSTM deep neural network. IFAC-PapersOnLine, 58(4), 79–84. https://doi.org/10.1016/j.ifacol.2024.07.197

[18]. Jun, S., Jung, D., & Lansey, K. (2024). Three-dimensional convolutional neural network for leak detection and localization in smart water distribution systems. Water Research X, 25, 100264. https://doi.org/10.1016/j.wroa.2024.100264

[19]. Farah, E., & Shahrour, I. (2026). Standardization and benchmarking of datasets and evaluation protocols for machine learning-based water leak detection: A systematic review. Water, 18(18), 2267. https://doi.org/10.3390/w18182267

[20]. Ranganathan, S., Vinothkumar, N., & Muthuramalingam, E. (2026). Benchmarking diverse learning architectures for leak detection and localization in water distribution systems. Journal of Water Process Engineering, 87, 110114. https://doi.org/10.1016/j.jwpe.2026.110114