Sudipkumar Ghanvat, 2026, 14:1
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Enhancing Water Network Resilience: A Generative AI Framework for Proactive Leakage and Loss Detection

Sudipkumar Ghanvat[1], Aditi Shintre[2], Sohail Hawaldar[3]
[1]Sr. Director & Head – Data & AI, VRIO Digital Dallas, USA
[2,3]Research Engineer, Neowesolutize Technology Pvt Ltd Pune

Abstract- Non-Revenue Water (NRW) is a major problem to urban water utilities across the globe, but physical leakage in old and complicated distribution systems is the main cause. Although sensing technologies, smart water networks, and data-based analytics have advanced, the majority of current leakage detection methods are reactive, depending on historical data, predetermined fault events, or discriminative models of machine learning with a limited capacity to predict rare or previously unknown leakage events. This review is a critical analysis of the development of leakage detection methodologies, including traditional methods of leakage detection such as physical method, artificial intelligence-based anomaly detection, and Digital Twin-based monitoring frameworks. The analysis shows that Digital Twins offer useful real-time system visibility and operational decision support, but their predictive functions are limited to the dependencies on scenarios and the lack of data. In order to overcome these constraints, this paper identifies the new role of Generative Artificial Intelligence as a game changer of proactive water network management. Adversarial and probabilistic generative models provide the capability to train the underlying distribution of multivariate time-series data, and to generate realistic and physically plausible leakage and anomaly scenarios. This feature can be used to augment data, train an anomaly detector, stress-test network behavior, and detect subtle or new leak signatures when incorporated into a physics-aware Digital Twin environment. The review summarizes the latest advances in the field of generative time-series modeling, anomaly detection, and Digital Twin integration, and comments on their applicability in industry, implementation issues, and considerations to implement them on a utility-scale basis. The main gaps in the research are defined, such as physics -informed generative models, explainable AI to gain the trust of operators, and field validation on a large scale. In general, the paper finds Generative AI-Enhanced Digital Twins as the prospect of predictive maintenance, enhanced network resilience, and sustainable NRW reduction in the future smart water system.

Keywords: Non-Revenue Water; Leak Detection; Generative Artificial Intelligence; Digital Twin; Smart Water Networks; Anomaly Detection; Predictive Maintenance.

I. INTRODUCTION

The urban water distribution systems are the foundation of the social health, economic activity and environmental sustainability. Nevertheless, these networks are confronted with consistent issues of aging infrastructure, urbanization, stress due to climate, and complexity of operations. Non-Revenue Water (NRW) is one of the most important and prevalent challenges facing water utilities around the globe and it is defined as the water that is generated and not charged because of physical losses, apparent losses and unauthorized usage. In most urban areas, especially in developing and rapidly developing countries, the NRW rates are still too high, often reaching 30-40% which leads to significant economic loss, energy loss, and the unnecessary burden on the limited water resources (Liemberger and Wyatt, 2019).

Physical leakage of pipes, joints, valves and service connections is the largest proportion of NRW. Leakage not only minimizes water supply, but it also compromises the stability of pressure, hastens the rate of pipes wearing out, consumes more energy to pump the water and exposes the risk of contamination due to intrusion. This is further complicated by the reality that most leaks start off small, latent failures that can last many years before they start showing or resulting in disastrous explosions. Early detection of such early-stage leaks in large and complex networks, many of which are intermittently used, is a significant technical and operational challenge.

The leakage management has traditionally relied on the traditional methods of leak detection that include acoustic surveys, step-testing, district metered area (DMA) analysis, and manual field inspection. Although these are useful in identifying known or suspected leakages, they are generally reactive, labor-intensive and episodic because they need specific campaigns instead of allowing network-wide surveillance. In addition, they do not perform well in dense urban areas, plastic pipes, low-pressure settings, and intermittently supplied systems that are becoming more common in most areas (Li et al., 2015).

The spread of sensors, Supervisory Control and Data Acquisition (SCADA) systems, and smart metering technologies has contributed to the development of smart water networks, which allow monitoring pressure, flow, acoustic signals, and consumption patterns continuously (Sharma et al., 2019). This data environment has triggered the use of machine learning and data-driven anomaly detection methods to identify leakage. Supervised, unsupervised and deep learning models have been shown to be more sensitive and automated than purely physical methods (Nimri et al., 2023). Nevertheless, even with these developments, the majority of current AI-based systems are limited due to their dependence on the use of historical data and discriminative learning paradigms. The nature of leakage events is sparse, non-uniform and contextual, causing imbalance in classes, a lack of coverage of failure modes, and a lack of extrapolation to new or changing leakage signatures.

In line with the AI-based analytics development, Digital Twin technology has become popular as a way of combining physical system models with real-time data streams. Digital Twins in water distribution networks are complex systems that are hydraulic models, asset data, sensor data, and logic, which are used to construct dynamic virtual representations of physical systems (Canivete, 2025). These platforms allow real time monitoring, scenario simulation and decision support to provide utilities with better situational awareness and operational insight. The majority of existing Digital Twin applications are however essentially descriptive or diagnostic, based on predefined fault conditions, threshold-based alarms, or residual analysis. They are generally not able to reason in advance but instead perform what-if simulations using known parameters, and not predictive reasoning about unknown or emerging failure modes.

This disjunction between information-based detection with a constraint on historical experience and Digital Twins with a constraint on pre-existing situations underscores a major weakness of the existing water network analytics. Indeed, to be proactive in leakage management, one needs systems capable of predicting realistic yet unknown anomalies, pre - equip detection models and operators with conditions not yet experienced in the field. To deal with this challenge, it is important to shift away solely to discriminative and reactive models to models that can learn and explore the underlying distribution of system behavior.

Generative Artificial Intelligence (AI) provides such a feature. Generative models, in contrast to discriminative models, attempt to learn the probability distribution of the data, rather than of observed patterns. This way they are able to generate new realistic samples that capture both the common operating conditions and the possible deviations. Generative models have the capability to capture time-dependent relationships, multi-variable relationships, and long-term relationships, which are critical in modeling pressure, flow, and acoustic behavior in water networks in time-series domains. The capability to create a variety of realistic anomaly scenarios offers an effective solution to the problem of data scarcity, class imbalance, and generalization constraints that impede the existing leakage detection syst ems (Duraivel et al., 2024).

The application of Generative AI into a Digital Twin architecture provides a synergistic architecture. The Digital Twin offers physical context, system constraints, and spatial structure, whereas Generative AI offers the capability to investigate a vast spectrum of likely operating and failure cases not just in the historical record. This type of integration allows generating synthetic, but physically consistent, leakage patterns that can be utilized to augment training data, stress-test detectors, and aid in proactive decision-making. Notably, this strategy transforms the Digital Twin into an active exploratory platform rather than a passive reflection of the system that can facilitate predictive maintenance and resilience-based management (Li et al., 2025).

This review aims to critically review the development of leakage detection approaches in water distribution systems and to establish Generative AI-Enhanced Digital Twins as a viable paradigm of proactive leakage and loss control. The paper will start with a review of traditional physical and data-driven methods of detecting leakages and then proceed with the developments in AI-based anomaly detection and Digital Twin applications. It subsequently summarizes the recent advances in generative modeling and explains how the methods can be systematically combined with Digital Twins to address the current limitations. Lastly, the review outlines research gaps, practical issues and future research directions required to transform this new paradigm into operational utility-scale solutions. It is hoped that by closing the gap between generative intelligence and physics-aware Digital Twins this work will be part of the current shift towards smarter, more resilient, and sustainability-oriented water infrastructure management.

II. LEAKAGE DETECTION IN WATER DISTRIBUTION NETWORKS: STATE OF THE ART

Water leakage detection in water distribution networks has been a key issue of concern in water utilities as it directly affects Non-Revenue Water, infrastructure deterioration, and operational expenses. A broad variety of physical, hydraulic and monitoring-based methods have been invented over the years to detect and locate leaks. Although these methods have evolved to a mature level, recent publications continuously point out their inabilities to detect leakage at an early stage and provide a reliable, network-wide detection. The traditional approach to leakage management is based on physical and hydraulic approaches. A systematic review of leakage detection practices, including acoustic inspection, step-testing, and minimum night flow analysis, was performed by Puust et al. (2010). They determined that acoustic correlators and noise loggers can be used to detect medium and large leaks, especially in metallic pipes, but have low sensitivity in plastic pipelines and high-noise urban locations. The authors determined that the methods are most effective to confirm and localize leakages, but not to detect them proactively, since in most cases, these methods need a high leakage level to be able to identify it with any degree of reliability.

The renewed interest in hydraulic transient analysis has occurred in the recent years when sensing and modeling capabilities were enhanced. Duan et al. (2020) presented an extensive review of transient flow modeling of urban water supply systems and showed how the propagation and reflection of pressure waves may be used to detect anomalies including leakages and faulty valves. Their discussion revealed that smaller leaks can theoretically be detected by transient-based methods than steady-state methods. They, however, also noted that sensitivity to boundary conditions, operational disturbances and incomplete knowledge about the system inhibit practical implementation to ensure robustness in the actual operating conditions.

The development of sensing technologies has resulted in more usage of SCADA-and IoT-enabled monitoring systems. Oberascher et al. (2022) overviewed the progress towards smart water cities and pointed to the increased utilization of distributed pressure, flow, and acoustic sensors to monitor the situation 24/7. Their study showed that sensor data with high frequencies enhances spatial awareness and allows detecting anomalies in almost real -time. However, the authors emphasized that sensing infrastructure is not sufficient to ensure the reliability of leak detection because without sophisticated analytical models, the interpretation of raw data is quite difficult.

The recent research has been on high-resolution sensing and data-centric. Rousso et al. (2023) reviewed the literature concerning smart water networks that use high-frequency pressure and acoustic sensors in the field of operation. Their results indicated that dense sensor arrays enhance the detection sensitivity, especially in transient events, however, at the cost of data complexity and noise. The paper found that traditional threshold-based or rule-based analytics cannot fully utilize high-frequency data, which commonly results in false alarms or false detections due to changing demand conditions. Control and system-identification perspective has also been used to review leak detection and localization techniques. Romero-Bena et al. (2023) compared strategies of leakage detection in the hydraulic modeling, statistical analysis, and data-driven paradigm. Their survey indicated that the majority of the currently used approaches are based on the deviations of the historically observed behavior and thus fail in the case of the leakage patterns that have never been seen before. The authors underlined that the lack of data and the unbalanced classes were the basic obstacles on the way to the enhancement of the detection reliability, especially in small or gradually evolving leaks.

It has been suggested to use model-based approaches to enhance robustness, by using system dynamics explicitly. Vrachimis et al. (2021) used a model invalidation framework to detect leaks, showing that the differences between measured and simulated pressur es could be used to reveal the presence and location of leaks. Although they perform well in controlled situations, their findings demonstrated that model uncertainty and calibration errors can play a big role in detecting real networks which supports the reliance of such techniques on high-fidelity hydraulic models.

These findings have been summarized and gaps that persist in recent reviews. In their review of the water leak detection techniques, Farah et al. (2024) concluded that, even with the modern technological advancements, most of the traditional and monitoring-based techniques are reactive in nature. Their research highlighted that early leakage is usually not detected because of the small hydraulic signature and that network-wide scalability is a very important issue especially in massive and looped distribution systems. One of the promising directions in the field of early detection has been suggested as high-frequency pressure monitoring. Meniconi et al. (2024) proved that the combination of low- and high-frequency pressure data enhances the sensitivity to minor hydraulic anomalies.

Their work demonstrated better detection of small disturbances, but also noted the challenge of separating true leaks and operational transients without sophisticated pattern recognition systems. On the whole, the state-of-the-art literature points to a common agreement that traditional physical, hydraulic, and monitoring-based leak detection methods are efficient in detecting severe failures but are still insufficient to detect incipient leaks and extrapolate in complicated network settings. The main bottleneck is no longer sensing capability, but the absence of predictive intelligence that is able to predict rare, evolving, or unseen leakage patterns. This has inspired the transition to more sophisticated Digital Twin models and Generative Artificial Intelligence, which would address the problem of data scarcity and improve proactive leakage control.

III. AI-BASED APPROACHES FOR LEAK DETECTION AND ANOMALY ANALYSIS

The introduction of artificial intelligence into monitoring water distribution has greatly enhanced the ability to identify leaks and bursts as opposed to the traditional hydraulic and rule-based methods. Through the use of pressure, flow, acoustic, and smart metering sensors, AI-based techniques will be used to detect anomalous trends that suggest leakage automatically. Recent literature has shown a definite transition of classical machine learning models to deep learning and sophisticated representation learning algorithms, as well as highlights the continued issues of data imbalance and generalization.

The initial research on AI-based leakage detection mainly used supervised machine learning models that were trained with labeled sensor data. Coelho et al. (2020) introduced one of the first practical benchmarking studies on the deployment of real-time sensors, comparing the performance of Random Forests, Decision Trees, Support Vector Machines, and Neural Networks on leak detection. They showed that supervised models are capable of high classification accuracy when they are trained on representative datasets.

Nevertheless, sensitivity to noise, sensor reliability, and changing operating conditions were also noted in the study, and the reliance of the supervised methods on the high-quality labeled data was pointed out. More recent literature has generalized this to anomaly detection in smart water metering systems. To overcome the issue of class imbalance, Kanyama et al. (2025) integrated ensemble learning with resampling methodology to enhance the detection of rare cases of anomalous consumption. Their results revealed that ensemble models are highly effective as compared to single classifiers in imbalanced environments. However, the authors determined that these methods are limited by the diversity of historical events and they might not be able to identify new or emerging leakage patterns.

Deep learning has become a leading paradigm of modeling the non-linear time and space properties of leakage signals. Zhou et al. (2019) proved that deep neural networks are able to locate burst events in terms of pressure data more accurately than traditional feature-based approaches. Their contribution demonstrated that deep models are able to implicitly acquire nonlinear hydraulic relations to allow more accurate localization of bursts in distribution networks. Based on this premise, Wang et al. (2020) suggested a deep learning pipeline to detect bursts in District Metered Areas, combining prediction, classification, and correction steps. Their model was successful in minimizing false alarms due to fluctuations in demand and operational interruptions. The study however, also observed that model performance reduces when used in network conditions which were not used in training. Recent studies have extended deep learning to acoustic and hybrid sensing platforms. Ma et al. (2025) presented two-stage temporal segmentation method based on convolutional neural networks and incremental learning in order to solve non-stationary acoustic signals. Their findings indicated that adaptive learning is an effective method of enhancing detection accuracy in noisy conditions as new data is being received. On the same note, Liu et al. (2024) used contrastive learning to acoustic leak detection and showed better ability to discriminate between leak-induced and non-leak vibroacoustic patterns with fewer false positives.

Another common problem with AI-based leak detectors is the extreme imbalance between regular operation data and leakages. The incidences of leak are infrequent, non-uniform and frequently not well documented resulting in non-leak dominated datasets. Kanyama et al. (2025) specifically solved this problem by resampling and ensemble methods and presented better recall of anomalous events. Their study, however, also highlighted that resampling does not add new leakage behaviors to the system, which constrains the ability to be robust in unseen conditions. The lack of data is also enhanced by the lack of spatial uniformity among networks. Yu et al. (2024) addressed this issue by converting the pressure time-series into spatial images through kriging interpolation and classifying them with the help of convolutional neural networks. Although this method has been found to have high true positive rates, its performance relies on sensor density and spatial data quality which might not be evenly distributed across actual networks.

Regardless of significant performance gains, discriminative AI models are still inherently limited by the fact that they are dependent on past data distributions. Leite et al. (2024) suggested an online burst detection algorithm with dynamic shape similarity metrics, which allows processing DMA flow data in near real-time. Their strategy enhanced responsiveness yet was still needed to be adapted to local operating conditions, which restricted transferability. Generalization is tried to be improved in spatial and hybrid modeling methods. Elshazly et al. (2024) presented a spatial machine learning model built on GIS, which incorporates geospatial characteristics with sensor information, which is more accurate in detection than a solely temporal model. Chen and Wang (2024) also established that acoustic sensing based on hydrophones and ensemble learning could be effectively used to produce strong performance with low false alarms in actual water networks. However, both papers recognized that discriminative models fail when they are faced with leakage patterns that are beyond their training scope. Altogether, the analyzed literature suggests that although AI-based solutions are highly effective in automating processes and improving the accuracy of detection, they are rather reactive. Their operation is highly dependent on the diversity and representativeness of historical data, which restricts the predictability of rare, emergent, or never-before-seen leakage situations. Such constraints drive the investigation of generative and hybrid modeling paradigms with the ability to learn the underlying data distribution and generate realistic new leakage patterns.

IV. DIGITAL TWINS FOR SMART WATER NETWORKS

The Digital Twin (DT) technology has become one of the enablers of the digital transformation of water distribution networks because it offers a constantly updated virtual image of physical resources, their working conditions, and dynamics of the system. In contrast to conventional off-line hydraulic models, Digital Twins are almost real-time and combine sensor data with simulation engines and analytics to aid in monitoring, diagnosis, and decision-making. Recent studies show that Digital Twins are a great way to improve situational awareness and operational control, but their efficiency in the early detection and prediction of leaks is limited due to inherent design constraints.

Initial applications of Digital Twin in water systems were mostly concerned with real -time monitoring and decision support. Brahmbhatt et al. (2023) designed a decision support system based on DT with the help of hydraulic simulation and real-time sensing on the WNTR framework. Their system allowed it to manage adaptive pressure and localize the leaks using residual analysis between simulated and observed pressure. The significant contribution of this work is the ability to show how Digital Twins can be transformed into a platform rather than a planning tool. Nevertheless, leak detection was based on preset thresholds and model residuals which restricted the sensitivity to thin or early leakages and the performance was determined by the accuracy of model calibration.

With the rise in complexity of Digital Twin, researchers started to consider the applicability of large-scale networks. Wu et al. (2023) suggested a high-fidelity Digital Twin model of smart water grids with thousands of kilometers of pipelines. Their solution was a combination of physics-simulation and data-driven analytics to identify abnormalities like bursts and unauthorized usage. One of the innovations of this work was its scalability and deployment-oriented testing, with lead times of detection of up to several days in some cases. Irrespective of this development, anomaly detection remained very reactive and relied on variations that surpassed learned thresholds, limiting the ability to detect slow changing leakage phenomena.

Recent research has examined Digital Twins that have been specifically developed to detect leakages. Singh et al. (2024) introduced a Digital Twin architecture of large-scale water distribution systems that combines hydraulic simulation with graph-based learning to detect leakage indicators. Their case study involving a university campus network showed that localization accuracy was better than when hydraulic standalone models were used. The originality of this work is that it introduces the awareness of network topology into the DT framework. Nevertheless, the system had to be trained on leakage patterns which had been historically observed, which restricted its capability to extrapolate to unknown failure modes.

A different body of work has been interested in the connection between Digital Twins and machine learning classifiers. Pandey et al. (2024) created a Digital Twin to detect and localize anomalies based on machine learning and applied a two-step framework: hydraulic residual modeling followed by the classification of anomalous patterns. Their performance demonstrated a higher detection speed and less false alarms than purely physics - based DTs. The main weakness, though, was that the effectiveness of the classifier was determined by the variety of labeled anomalies to be used in the training process, which again underscored the need to use historical data. In addition to leakage detection, Digital Twins have been used more in system resilience and recovery planning. Dui et al. (2025) suggested a DT-based framework of the urban water distribution network resilience assessment that measures the significance of each pipe and maximizes recovery mechanisms following the failures. This work is novel since it goes beyond detection with Digital Twins to strategic resilience planning. However, the framework is oriented on the post-event analysis and recovery as opposed to the early-stage leak prediction.

Digital Twin environments have also attracted attention on predictive analytics. Syed et al. (2024) combined a Digital Twin and multimodal transformer networks to predict the usage of water and identify abnormalities. Their model showed a better prediction accuracy, which was a combination of physical simulation and sophisticated deep learning. Although this is a move towards predictive Digital Twins, the methodology is discriminative, as it only learns based on patterns that have been observed before and needs a lot of history to train. Flexibility and long-term development of water networks raise other issues. The DiTEC Digital Twin, proposed by Degeler et al. (2024), can manage the evolutionary changes of the network, including pipe aging and the network expansion. They used the concept of graph-based learning and model updates to overcome the static character of the traditional DTs. However, though useful in the case of slow structural changes, the method does not directly deal with the production or prediction of rare leakage cases.

Other recent literature has focused on the real-time detection of leakages in pressurized systems. Travaaz (2025) created a Digital Twin that has hardware sensing to detect and locate leaks in pipelines in real-time. This was shown to be practically feasible and able to detect early under controlled conditions. However, the logic of detection was still rooted in visualized deviations and not forecasted synthesis of possible failure modes. These trends are also validated by comprehensive survey of Digital Twin applications in the water industry. Ghorbani Bam et al. (2025) have reviewed the Digital Twin deployments systematically and found that the majority of the DTs are effective in visualization, monitoring, and scenario testing but do not have an autonomous intelligence to identify unknown failure modes.

On the same note, larger studies of digital transformation by Homaei et al. (2024) highlighted that although Digital Twins are the core of smart water systems, their analytical intelligence tends to be superficial in the absence of sophisticated learning layers. Recent Digital Twin–based approaches for water distribution networks are comparatively summarized in Table 1, highlighting their methodological focus, primary capabilities, and inherent limitations. Although Digital Twins have a substantial effect on real-time monitoring and operational decision-making, their analytical intelligence is most often limited to the predefined rules or discriminative learning models based on past data. These constraints limit their capability to predict uncommon, arising, or unobservable leakage conditions, highlighting the necessity of generative intelligence in Digital Twin models.

Table 1: Comparison of Digital Twin Approaches for Water Distribution Networks

Study	Digital Twin Approach	Primary Capability	Key Limitation
Brahmbhatt et al. (2023)	Hydraulic DT with real-time data assimilation and residual analysis	Leak localization and operational decision support	Threshold-based detection; sensitive to model calibration and demand uncertainty

Wu et al. (2023)	High-fidelity physics-based DT with data-driven analytics	Large-scale anomaly detection in urban networks	Reactive detection; limited sensitivity to small or gradual leaks
Singh et al. (2024)	DT integrating hydraulic simulation with graph-based learning	Improved spatial leak localization	Requires historical leak patterns; limited generalization to unseen failures
Pandey et al. (2024)	Hybrid DT combining simulation residuals with ML classification	Faster anomaly detection with reduced false alarms	Performance depends on labeled anomaly data
Syed et al. (2024)	DT integratedwith Multimodal transformer networks	Predictive usage and anomaly detection	Discriminative learning; requires extensive historical data
Degeler et al. (2024)	Adaptive DT for network evolution (DiTEC)	Handles long-term structural changes	Does not anticipate rare ornovel leakage scenarios
Travaš (2025)	Hardware-integrated real-time DT	Real-time leak detection in pressurized pipelines	Detection relies on observable deviations; limited predictive intelligence
Ghorbani Bam et al. (2025)	Survey of DT architectures in water sector	Comprehensive synthesis of DT applications	Identifies lack of autonomous predictive capability in current DTs

V. GENERATIVE AI–ENHANCED DIGITAL TWIN FRAMEWORK FOR PROACTIVE LEAK DETECTION

The latest developments in the field of artificial intelligence and Digital Twin technologies have led to a great enhancement in the monitoring and operational management of water distribution networks. Nevertheless, as noted in the above sections, the current systems are still largely reactive and are based on discriminative models and predetermined situations that are not much able to generalize beyond the previously observed leakage trends. This weakness drives the adoption of Generative Artificial Intelligence (AI) in Digital Twin settings to support proactive leakage detection, classification, and predictive maintenance.

The theory behind the proposed framework is the generative modeling theory. The seminal article of Goodfellow et al. (2014) proposed Generative Adversarial Networks (GANs), which showed that adversarial learning allows models to learn complicated data distributions without having explicit physical models. This observation defined that only observed samples could be used to produce realistic synthetic data. Simultaneously, Variational Autoencoders (VAEs) were proposed by Kingma and Welling (2013), which is a probabilistic generative model that models uncertainty by means of latent variables. Collectively, these works give the theoretical framework to generalizing the plausible system behaviors that cannot be determined by available information, which is a crucial step towards predicting the occurrence of rare leakage events in water networks.

Water distribution systems produce very organized multivariate time-series data that is affected by cycles of demand, hydraulic interactions, and operational controls. Having identified this, Yoon et al. (2019) have introduced TimeGAN, a hybrid model that incorporates adversarial learning and sequence modeling to maintain the temporal relationships and cross -variable associations. Their findings showed that they were more faithful to synthetic time-series generation than traditional GANs and VAEs. This ability can be directly applied to the modeling of pressure, flow and acoustic signals, where leakage signals develop over time and are frequently obscured by normal variability of operations.

Esteban et al. (2017) proposed further refinement of generative time-series modeling with recurrent conditional GANs. They used their method to produce time-series with desired characteristics by conditioning sequence generation on contextual variables. When applied to leakage detection, this conditioning enables the creation of a variety of leakage scenarios as a parameter of magnitude, position, or rate of onset, and offers a systematic way to explore scenarios in a Digital Twin.

The real-world water networks are typified by noisy, incomplete and irregular sensor data. To deal with this issue, Fortuin et al. (2020) suggested the GP-VAE framework, which combines Gaussian process priors and variational autoencoders to enhance resistance to missing data and noise. Their article showed that probabilistic generative models are more reliable in engineering use, which supports the appropriateness of VAEs in sensor-based Digital Twin systems.

In addition to data synthesis, generative models have demonstrated a great potential in unsupervised anomaly detection. Schlegl et al. (2017) showed that GANs trained on normal data only are capable of detecting anomalies through reconstruction error, which is an effective way of detecting deviations to the learned normal behavior. The paradigm is especially applicable to water networks, where the amount of labeled leak data is limited. This method was further enhanced by Zenati et al. (2018) who proposed efficient GAN-based anomaly detection with explicit inverse mappings, which makes the method more stable and computationally efficient. All of these works demonstrate that generative models can be useful in identifying subtle, previously unnoticed anomalies.

Water distribution systems leakage is usually observed as coordinated sensor channel changes. Li et al. (2019) introduced MAD-GAN, a multivariate anomaly detector model, to model the time-series variables correlation explicitly, to capture such dependencies. Their results showed that they can perform better than univariate methods in detection, which is the significance of multivariate generative modeling in complex infrastructure systems.

Although generative models offer strong data-driven functionality, unconditional generation can produce physically unrealistic situations. In this regard, Karpatne et al. (2017) and Willard et al. (2022) focused on the combination of scientific knowledge with machine learning models. They demonstrated in their work that physics-informed constraints enhance generalization, interpretability, and trustworthiness in scientific and engineering applications. Water distribution networks in particular can benefit from these principles, as hydraulic laws can be used to control and direct generative processes.

In the suggested model, this physical context is provided by the Digital Twin. Digital Twins as Rasheed et al. (2020) and Fuller et al. (2020) explain are aligned virtual models of real systems but need more sophisticated AI to go beyond description and diagnostics. The framework can be trained to generate a physics-aware generative model within the Digital Twin to allow exploration of plausible leakage scenarios by continuously updating the model under hydraulic feasibility and network topology constraints.

The generative model is operationally trained on historical data that reflects the normal behavior of the system, and then applied to generate a rich library of leakage and anomaly scenarios. These artificial patterns enrich actual data and are utilized to educate powerful anomaly detectors in the Digital Twin. In real-time mode, the incoming sensor data is compared with historical and synthetic signatures, which allows early detection, classification and localization of leaks. This preventive maintenance feature helps to implement a shift in the reactive repair approach to predictive maintenance, where the intervention is informed by the early warning signs and risk-based prioritization.

The analyzed literature shows clearly that Generative AI can be the missing intelligence layer that will allow overcoming the drawbacks of the present Digital Twin-based leakage detection systems. Generative models allow Digital Twins to become more predictive, proactive systems by learning the underlying distribution of system behavior and synthesizing realistic, physically constrained leakage scenarios. This combined framework is the foundation of resilient and intelligent water network management that will help to significantly decrease non-revenue water and increase the sustainability of infrastructure in the long term.

VI. INDUSTRIAL RELEVANCE, CHALLENGES, AND PRACTICAL CONSIDERATIONS

Utility-scale applicability and operational benefits

Non-Revenue Water (NRW) reduction is a technical objective at utility scale, but a quantifiable tool to enhance service coverage, reliability, and financial sustainability. This operational framing is supported by the evidence in the Indian context. Based on the 55 Indian metropolises quantitative efficiency analysis, Bandari and Sadhukhan (2023) demonstrate that the degree of NRW management is strongly correlated with the performance outcomes in terms of per-capita supply and cost recovery, which supports the idea that NRW reduction can provide material system-level benefits without necessarily raising the availability of raw water.

In the context of industrial operations, these benefits can be translated into practical benefits: the earlier leakage is detected, the less time is wasted on production and pumping, the more stable the pressure, and utilities can focus on repairs, which will save time on searching and minimise the cost of excavations. Nevertheless, the above advantages can only be achieved when the outputs of analytics are integrated into workflows (alarm triage, dispatch, repair verification) and complemented by regular data governance.

Conceptual relevance for global utilities such as Veolia

The conceptual placement of Generative AI-enhanced Digital Twins is in line with the current way large operators organize digital water services. Publicly, Veolia presents Hubgrade as a set of digital solutions that integrate data, AI, and human control to monitor and optimize business in environmental services, such as water. Their press release on the global implementation of an AI-based Hubgrade solution and their Hubgrade digital services pages depict the direction of the industry in terms of centralized monitoring, alarm management, and optimization, which is the exact environment in which an AI-enhanced Digital Twin can be implemented as a decision-support layer. That is, the main industrial use of your framework is not another model, but a capability upgrade: generative scenario synthesis to stress-test, train, and harden detectors such that utilities can detect small leaks sooner and respond to previously unknown signatures of leaks with fewer false alarms.

Key challenges: model fidelity, computational cost, and data quality

One of the first-order deployment constraints is model fidelity. The usefulness of Digital Twins lies in the fact that they capture physical structure and limits, but the accuracy of their implementation is limited by calibration, demand uncertainty, and sensor representativeness. The DT literature stresses that Digital Twins can be useful when modeling decisions and data integration are approached thoughtfully, and that enablers and challenges (such as model quality and complexity of integration) have a very strong influence on real-life impact.

In the case of Generative AI, the concept of fidelity is even more of a concern: synthetic anomalies should be physically realistic or a detector can get immune to simulated anomalies instead of real leaks. That is why physics-aware ML principles, which are explicitly suggested in scientific/engineering AI literature, are significant when integrating generative models with Digital Twins (e.g. limiting generation by hydraulic feasibility and network topology).

The operationally important factors are also computational cost and latency. GAN/time-series generative models are potentially costly to train, and might need centralized compute, whereas alarm detection needs to be available in near real time. This is often understood in practice in an architecture where a heavy generative training and scenario library update is done offline (periodically), and lightweight anomaly detectors are continuously run in production.

The most difficult obstacle in live water networks is usually data quality. Even normal data may be polluted by latent leaks, temporary supply regimes, sensor drift and missing values. The evidence of the Indian governance on digitalization highlights that the advantages of SCADA/smart metering are highly dependent on the quality of implementation and the capacity of institutions- not only the installation of technology. Banerjee et al. (2022) compare the case of Bengaluru and Singapore and examine such digital technologies as smart water metering and SCADA, and highlight that to make them work, sequencing, integration, and governance are needed instead of the adoption of individual tools. In your framework, it implies that you should train the Generative AI layer on carefully selected clean normal periods with explicit handling of missing data and drift otherwise synthetic generation will recreate the noise and bias that you do not want.

Interpretability, trust, and real-world deployment constraints

Interpretability and trust are critical to adoption: operators must understand why an alert has been raised, its level of confidence, and what to do next. It is here that the Digital Twin context can prove practically valuable: DT-based reasoning can be used to give an operational explanation of the residual patterns in pressure (such as leak-like demand increase in this zone), and can be used to do counterfactual checks (such as whether the anomaly is similar to an operation of a known valve). In the meantime, generative models can enhance transparency by recalling the nearest synthetic scenario prototiles (this event is similar to a gradual leak onset pattern) and measure uncertainty. Such explanations in industrial contexts such as Hubgrade - style monitoring minimize alarm fatigue and raise the chances of analytics being taken action on.

The common deployment constraints are cybersecurity controls, integration with the legacy SCADA/CMMS systems, workforce preparedness, and must have quantifiable KPIs (detection lead time, false alarm rate, repair confirmation time). In the case of Indian utilities, other practical limitations frequently involve intermittent supply, patchy sensor coverage, and lack of labelled leak ground truth, and it is therefore necessary to have a strong validation and rollout (pilot DMA to city zones to full network) strategy.

VII. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

This review has discussed the latest developments in the field of leakage and loss detection in the context of physical inspection, data-driven analytics, and Digital Twin-based monitoring of water distribution networks. As demonstrated in the literature, although conventional and AI-assisted methods have enhanced the speed and coverage of detection, the majority of the existing methods are reactive, and lack a fundamental ability to generalize to rare or unseen leakage events. Digital twins have enhanced the visibility of the systems and the operational decision-making, although their predictive ability is typically limited to defined situations instead of open-ended risk anticipation.

In this context, Generative Artificial Intelligence is a new potential of increasing the resilience of water networks. Generative models can be used to create realistic leakage and anomaly conditions that are not limited to observed data by learning the underlying distribution of the normal system behavior. This ability, when incorporated into a Digital Twin environment, assists in proactive detection, classification, and localization of leaks, allowing utilities to plan ahead of the emergent or subtle failures before they deteriorate. With such integration, there will be a transition to predictive maintenance rather than the reactive repair practices, better allocation of resources, improved operational efficiency and sustainability of infrastructure in the long term.

However, there are a number of challenges that are open to it. It is also important to ensure that the synthetic scenario generation is physically faithful because anomalies that are unrealistic can compromise the reliability of the detector. Sensitivity to noise, missing values, and polluted normal baselines are still a problem to both the generative and discriminative models due to data quality concerns. Also, interpretability and operator trust are required in the real world, actionable explanations and confidence-sensitive alerts are needed to make sure that advanced analytics become useful components of utility processes.

Future studies are to create physics-informed generative models that can be trained on hydraulic constraints, create explainable generative AI methods that are specific to operational decision-making and perform large-scale pilot studies to confirm performance in real network conditions. It should also focus on deployment-oriented design, such as computational efficiency, compatibility with current monitoring systems, and standardized evaluation metrics. It will be important to address these directions to translate Generative AI-Enhanced Digital Twins out of the conceptual frames into the practical tools of resilient and sustainable water network management.

REFERENCES

1. Liemberger, R., & Wyatt, A. (2019). Quantifying the global non-revenue water problem. Water Supply, 19(3), 831-837.
2. Li, R., Huang, H., Xin, K., & Tao, T. (2015). A review of methods for burst/leakage detection and location in water distribution systems. Water Science and Technology: Water Supply, 15(3), 429-441.
3. Sharma, N., & Ramachandran, R. K. (2019, November). Security challenges for water distribution system using supervisory control and data acquisition (scada). In 2019 Fifth International Conference on Image Information Processing (ICIIP) (pp. 234-239). IEEE.
4. Nimri, W., Wang, Y., Zhang, Z., Deng, C., & Sellstrom, K. (2023). Data-driven approaches and model-based methods for detecting and locating leaks in water distribution systems: Aliterature review. Neural Computing and Applications, 35(16), 11611-11623.
5. Canivete, C. R. C., Chitauro, M., Flörke, M., & Okorie, M. E. (2025). A Bibliometric and Network Analysis of Digital Twins and BIM in Water Distribution Systems. Technologies, 13(12), 575.
6. Duraivel, S., Gopal, V., & Kannan, P. (2024). Using Generative Adversarial Networks (GANs) for Predictive Water Management and Anomaly Detection in Smart Water Systems to Achieve SDG 6. Available at SSRN 5057104.
7. Li, T., Long, Q., Chai, H., Zhang, S., Jiang, F., Liu, H., ... & Li, Y. (2025). Generative ai empowered network digital twins: Architecture, technologies, and applications. ACM Computing Surveys, 57(6), 1-43.
8. 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.
9. Duan, H. F., Pan, B., Wang, M., Chen, L., Zheng, F., & Zhang, Y. (2020). State-of-the-art review on the transient flow modeling and utilization for urban water supply system (UWSS) management. Journal of Water Supply: Research and Technology—AQUA, 69(8), 858-893.
10. Bakhtawar, B., & Zayed, T. (2021). Review of water leak detection and localization methods through hydrophone technology. Journal of Pipeline Systems Engineering and Practice, 12(4), 03121002.
11. 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.
12. Oberascher, M., Rauch, W., & Sitzenfrei, R. (2022). Towards a smart water city: A comprehensive review of applications, data requirements, and communication technologies for integrated management. Sustainable Cities and Society, 76, 103442.
13. Romero-Ben, L., Alves, D., Blesa, J., Cembrano, G., Puig, V., & Duviella, E. (2023). Leak detection and localization in water distribution networks: Review and perspective. Annual Reviews in Control, 55, 392-419.
14. Vrachimis, S. G., Timotheou, S., Eliades, D. G., & Polycarpou, M. M. (2021). Leakage detection and localization in water distribution systems: A model invalidation approach. Control Engineering Practice, 110, 104755.
15. Farah, E., & Shahrour, I. (2024). Water leak detection: a comprehensive review of methods, challenges, and future directions. Water, 16(20), 2975.
16. Meniconi, S., Rubin, A., Tirello, L., Doro, A., Brunone, B., & Capponi, C. (2024). Mapping pressure surge source in urban water networks:

Integrating low-and high-Frequency pressure data with an illustrative real case study. Water Resources Research, 60(8), e2023WR036773.
17. Zhou, X., Tang, Z., Xu, W., Meng, F., et al. (2019). Deep learning identifies accurate burst locations in water distribution networks. Water Research, 166, 115058.
18. Wang, X., Xu, K., & Smith, K. (2020). Burst detection in district metering areas using deep learning method. Journal of Water Resources Planning and Management, 146(6).
19. Coelho, J. A., Glória, A., & Sebastião, P. (2020). Precise water leak detection using machine learning and real-time sensor data. IoT, 1(2), 474–493.
20. Leite, R., Amado, C., & Azeitona, M. (2024). Online burst detection in water distribution networks based on dynamic shape similarity measure. Expert Systems with Applications.
21. Kanyama, M. N., Bhunu Shava, F., Gamundani, A. M., & Hartmann, A. (2025). AI-driven anomaly detection in smart water metering systems using ensemble learning. Water, 17(13), 1933.
22. Ma, X., et al. (2025). Leak detection in water supply networks using two-stage temporal segmentation and incremental learning. Water Research X, 100333.
23. Yu, H., et al. (2024). Leak detection in water distribution networks based on deep learning and kriging interpolation method. Aqua, 73(8), 1741–1755.
24. Liu, R., et al. (2024). Contrastive learning method for leak detection in water distribution systems. npj Water Research & Technology.
25. Chen, W., & Wang, H. (2024). AIoT-driven acoustic leak detection in real water networks using ensemble machine learning. Water Resources Management.
26. Elshazly, D., et al. (2024). Automated GIS-based spatial machine learning method for leak detection in water distribution networks. Applied Sciences, 14(13), 5853.
27. Brahmbhatt, P., Maheshwari, A., & Gudi, R. D. (2023). Digital twin assisted decision support system for quality regulation and leak localization in water distribution networks.
28. Wu, Z. Y., et al. (2023). High-fidelity digital twin–based anomaly detection and localization in smart water grids.
29. Singh, A., Maheshwari, A., & Singh, S. (2024). Digital Twin framework for leakages detection in large-scale water distribution systems. IFAC
30. Pandey, P., et al. (2024). Machine learning–based Digital Twin for water distribution network anomaly detection and localization. Engineering Proceedings.
31. Dui, H., Cao, T., & Wang, F. (2025). Digital Twin–based resilience evaluation and intelligent strategies of smart urban water distribution networks.
32. Syed, T. A., et al. (2024). Predictive usage and leakage detection using Digital Twin and multimodal transformer networks. Water.
33. Degeler, V., et al. (2024). DiTEC: Digital Twin for evolutionary changes in water distribution networks. Springer.
34. Travaš, V. (2025). Digital twin for real-time leakage detection and location in pressurized pipelines. Journal of Hydroinformatics.
35. Ghorbani Bam, P., et al. (2025). Digital Twin applications in the water sector: A review. Water.
36. Homaei, M. H., et al. (2024). Digital transformation in water distribution systems based on the Digital Twins concept.
37. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.
38. Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.
39. Yoon, J., Jarrett, D., & Van der Schaar, M. (2019). Time-series generative adversarial networks. Advances in neural information processing systems, 32.
40. Esteban, C., Hyland, S. L., & Rätsch, G. (2017). Real-valued (medical) time series generation with recurrent conditional gans. arXiv preprint arXiv:1706.02633.
41. Fortuin, V., Baranchuk, D., Rätsch, G., & Mandt, S. (2020, June). Gp-vae: Deep probabilistic time series imputation. In International conference on artificial intelligence and statistics (pp. 1651-1661). PMLR.

42. Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., & Langs, G. (2017). Unsupervised anomaly detection with generative adversarial networks. In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI).
43. Zenati, H., Foo, C. S., Lecouat, B., Manek, G., & Chandrasekhar, V. R. (2018). Efficient gan-based anomaly detection. arXiv preprint arXiv:1802.06222.
44. Li, D., Chen, D., Jin, B., Shi, L., Goh, J., & Ng, S. K. (2019, September). MAD-GAN:
45. Multivariate anomaly detection for time series data with generative adversarial networks. In International conference on artificial neural networks (pp. 703-716). Cham: Springer International Publishing.
46. Willard, J., Jia, X., Xu, S., Steinbach, M., & Kumar, V. (2022). Integrating scientific knowledge with machine learning for engineering and environmental systems. ACM Computing Surveys, 55(4), 1-37.
47. Karpatne, A., Watkins, W., Read, J., & Kumar, V. (2017, December). How can physics inform deep learning methods in scientific problems?: Recent progress and future prospects. In 31st Conference on Neural Information Processing Systems (NeurIPS) (pp. 1-5).
48. Rasheed, A., San, O., & Kvamsdal, T. (2020). Digital twin: Values, challenges and enablers from a modeling perspective. IEEE access, 8, 21980-22012.
49. Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: enabling technologies, challenges and open research. IEEE access, 8, 108952-108971.
50. Bandari, A., & Sadhukhan, S. (2023). Efficiency of non-revenue water reduction in improving water supply performance in Indian metropolises. Water Supply, 23(5), 1917. =
51. Banerjee, C., et al. (2022). Digitalization in Urban Water Governance: Case Study of Bengaluru and Singapore. Frontiers in Environmental Science. (PDF) https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2022.816824/pdf
52. Veolia (Press release, 2023). Worldwide deployment of Hubgrade digital solutions (AI-based monitoring).https://www.veolia.com/en/our-media/press-releases/veolia-worldwide-deployment-digital-solution-based-artificial-water-energy-waste
53. Hubgrade Digital Services (Veolia Water Technologies). Hubgrade / digital twin modules and AI-enabled monitoring overview. https://www.veoliawatertech.com/en/solutions/services/hubgrade-digital-services