



Explainable AI (XAI) Digital Twin Framework for Predictive Failure Detection and Lifecycle Management of U.S. Drinking Water Treatment Infrastructure

Taufiqur Rahaman;

[1]. Master in Civil and Environmental Engineering, College of Engineering, Lamar University, Texas, USA;
Email: rueconomy@gmail.com

Doi: [10.63125/v6bmmt68](https://doi.org/10.63125/v6bmmt68)

Received: 19 January 2026; Revised: 24 February 2026; Accepted: 19 March 2026; Published: 06 June 2026;

Abstract

This study examined the problem of opaque, reactive, and fragmented condition monitoring in U.S. drinking water treatment infrastructure, where pumps, membranes, filtration units, disinfection systems, and distribution assets are frequently maintained on fixed schedules or after failure, and where data-driven predictive models, when used at all, often function as black boxes that operators cannot interpret or trust. The purpose of the study was to assess how an Explainable AI (XAI) Digital Twin Framework, integrating physics-informed digital twin modeling with interpretable machine-learning methods, influences predictive failure detection and lifecycle management performance for drinking water treatment assets. A quantitative, cross-sectional, case-based design was adopted, and data were collected through a structured five-point Likert-scale questionnaire from 138 valid respondents out of 155 distributed questionnaires, representing an 89.0% valid response rate. The sample included water treatment operators, control and instrumentation engineers, data scientists and machine-learning engineers, asset-management and maintenance planners, and utility managers and regulatory staff, with 65.9% directly involved in condition-monitoring or asset-management activities. The key variables were digital twin modeling and integration, XAI model interpretability, sensor data quality and validation, predictive failure detection accuracy, operator and institutional engagement, framework design quality, and lifecycle management performance. The analysis plan included descriptive statistics, reliability testing using Cronbach's alpha, Pearson correlation, regression modeling, a framework maturity index, and an asset risk-control priority matrix. The headline findings showed that all major constructs were rated high, with lifecycle management performance recording the highest mean score of 4.19, followed by digital twin modeling and integration at 4.14 and framework design quality at 4.10. Reliability was acceptable to excellent, with Cronbach's alpha values ranging from 0.80 to 0.93. Correlation results showed significant positive relationships, including $r = 0.77$ between framework design quality and lifecycle management performance. Regression results confirmed that the model explained 70.2% of the variance in lifecycle management performance, $R^2 = 0.702$, adjusted $R^2 = 0.689$, $F(6,131) = 51.36$, $p < 0.001$. Framework design quality was the strongest predictor, $\beta = 0.33$, followed by digital twin modeling and integration, $\beta = 0.26$, and XAI model interpretability, $\beta = 0.23$. The findings imply that U.S. water utilities should strengthen physics-informed digital twin modeling, model explainability, sensor validation, and operator engagement to improve the reliability and transparency of predictive asset management.

Keywords

Explainable AI; Digital twin; Predictive maintenance; Drinking water infrastructure; Lifecycle asset management.

INTRODUCTION

Predictive failure detection refers to the process of anticipating the degradation or breakdown of engineered assets before failure occurs, using data-driven models that learn from historical and real-time condition information. In U.S. drinking water treatment infrastructure, predictive failure detection is closely associated with the health of pumps, membranes, filtration units, disinfection systems, sensors, and distribution assets, whose unexpected failure can interrupt supply, compromise water quality, and threaten public health. A digital twin in this context is a dynamic, physics-informed virtual representation of a physical asset or system that is continuously updated with sensor data, enabling analysts to simulate behavior, test scenarios, and forecast condition (Grieves & Vickers, 2017). When such a twin is paired with Explainable AI (XAI), the predictions it generates are accompanied by human-interpretable reasons, so that operators can understand why a component is flagged as at risk rather than receiving an opaque score (Barredo Arrieta et al., 2020). In practical terms, an XAI digital twin framework treats predictive maintenance not as a black-box scoring exercise but as a transparent, auditable, and lifecycle-based process, so that model outputs, physical mechanisms, and operator judgment can be reconciled rather than kept separate.

Formally, the digital twin maintains an estimate of the asset state that is corrected as new sensor observations arrive. Let the estimated state at time t be updated from the previous state through a physics-based transition model f , driven by control inputs u and parameters θ , and corrected by the discrepancy between the measured observation z and the predicted observation h , weighted by a gain K . This data-assimilation update, given in Equation (1), is the mechanism by which the twin remains synchronized with the physical asset:

$$\hat{x}_t = f(\hat{x}_{t-1}, u_t; \theta) + K_t(z_t - h(\hat{x}_t^-)) \quad (1)$$

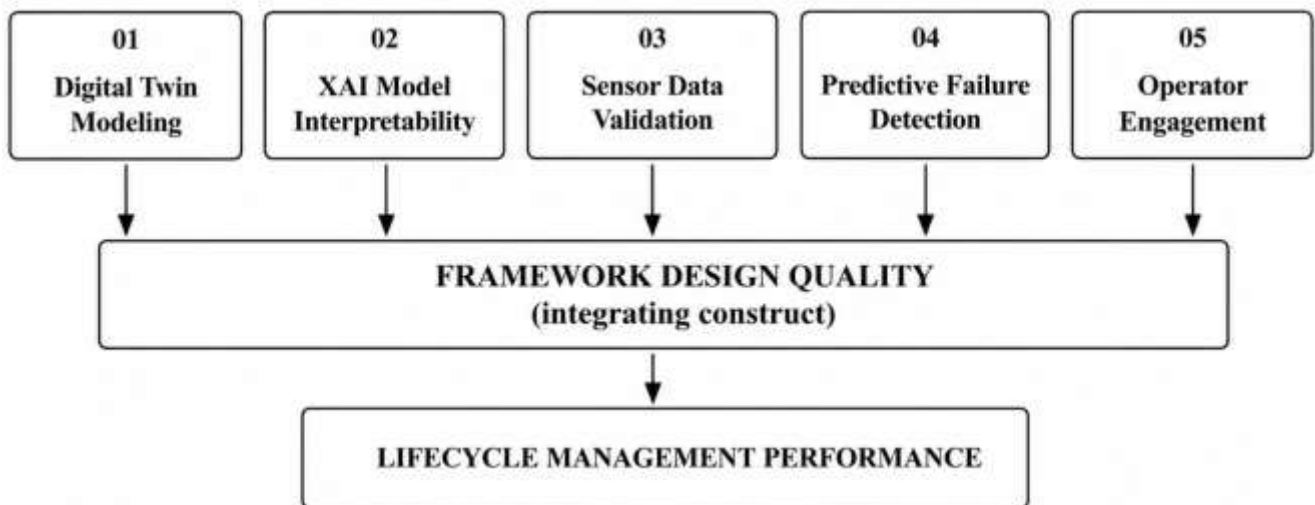
Here, $\hat{x}_t$ is the corrected (posterior) state estimate, $\hat{x}_t^-$ is the prior estimate produced by the transition model before the measurement is incorporated, and the gain K_t governs how strongly new data pull the twin toward the observation. A well-calibrated gain balances model fidelity against sensor noise, which is why the quality of the twin depends jointly on the physical model f and on the reliability of the incoming measurements z .

Within this study, the XAI Digital Twin Framework is defined as the systematic integration of physics-informed digital twin modeling, interpretable machine-learning methods, sensor data validation, predictive failure detection, and operator engagement mechanisms that together support the condition monitoring and lifecycle management of drinking water treatment assets (Abdur & Iftikhar, 2021). Lifecycle management performance is defined as the assessed and measurable ability of the framework to detect emerging failures early, characterize asset condition accurately, prioritize interventions correctly, extend asset service life, and sustain auditable, repeatable maintenance decisions over time. Digital twins provide continuous virtual mirrors of physical systems by assimilating operational data, which makes them well suited to representing the dynamic behavior of treatment trains, pumping stations, and distribution networks (Tao et al., 2019). XAI provides the interpretive layer in which model predictions are explained through feature attributions, surrogate models, or inherently interpretable structures, so that utility staff and regulators can scrutinize and trust the outputs (Ribeiro et al., 2016).

The national significance of an XAI digital twin approach is connected to the scale, age, and criticality of U.S. drinking water infrastructure, much of which was installed decades ago and is approaching or exceeding its design life (ASCE, 2021). Utilities face large and recurring investment needs to repair, rehabilitate, and replace aging assets, and they must make these decisions under budget constraints and growing regulatory expectations (U.S. EPA, 2020). Pipe and equipment failures are frequent and costly, and their consequences extend beyond direct repair to service interruptions, water loss, and public-health risk (Folkman, 2018). Predictive, condition-based approaches offer a way to move from reactive or fixed-schedule maintenance toward interventions timed to actual asset condition, but their adoption in the water sector has been constrained by data limitations, integration challenges, and, critically, a lack of transparency in the models themselves. Because water utilities operate under strong accountability requirements, models whose reasoning cannot be explained are difficult to adopt, justify

to regulators, or defend to the public (Rudin, 2019).

Figure 1: XAI Digital Twin Framework for Predictive Failure Detection and Lifecycle Management



Digital twins and XAI provide the central methodological foundation for this research because together they frame predictive maintenance as a transparent, lifecycle activity that begins with asset and data characterization and continues through modeling, explanation, validation, interpretation, decision-making, and periodic updating (Rasheed et al., 2020). The combined approach requires utility staff to move from reactive repair toward anticipatory, evidence-based intervention, while retaining the ability to interrogate model reasoning (Hasan & Uddin, 2022; Mohiul & Badrul, 2022). Studies of digital twins in engineering repeatedly show that their value depends on both model fidelity and the quality of the data assimilated into them, and that a twin is only as reliable as the sensing and validation that feed it (Jones et al., 2020). Concerns about the interpretability of machine-learning models are especially important for water utilities because a prediction can appear accurate on historical data while remaining impossible to explain, audit, or generalize to new operating conditions, which undermines operator trust and regulatory acceptance (Adadi & Berrada, 2018).

A major technical issue in XAI digital twin frameworks is the reconciliation of predictive accuracy with interpretability, since the most accurate models are often the least transparent and the most transparent models are sometimes less accurate (Guidotti et al., 2018). Physics-informed approaches offer one route to reconciling the two by embedding known physical laws into data-driven models, which can improve both generalization and interpretability (Raissi et al., 2019). The relationship between digital twin modeling quality, model interpretability, and the resulting lifecycle management outcome is therefore central to this study, which treats digital twin modeling, XAI interpretability, sensor data validation, predictive detection accuracy, and operator engagement as independent variables because the degree to which a framework applies rigorous modeling, explanation, validation, and stakeholder involvement may influence framework design quality and the resulting lifecycle management performance.

Background of the Study

The background of this study is rooted in the condition of U.S. drinking water infrastructure and the growing gap between the demands placed on it and the tools available to manage it. Much of the nation's treatment and distribution infrastructure was constructed in the mid-twentieth century and is now aging, with many components operating well beyond their intended service life (ASCE, 2021). Utilities must manage this aging stock while meeting stringent water-quality standards, controlling costs, and responding to climate and demand pressures. Traditional maintenance in the sector has been largely reactive or based on fixed schedules, which can lead either to premature replacement of assets that still have useful life or to unexpected failures of assets that degrade faster than expected (Marlow et al., 2010).

Data-driven condition monitoring has emerged as a means of closing this gap. The proliferation of sensors, supervisory control and data acquisition systems, and advanced metering has made it feasible to observe the condition of treatment and distribution assets in far greater detail than before (Eggimann et al., 2017). Machine-learning methods can learn patterns of degradation from these data and forecast failures (Ran et al., 2019), and digital twins can integrate such forecasts with physical models to simulate asset behavior under different conditions (Pesantez et al., 2022). This combination is particularly valuable in water treatment, where the interaction of hydraulics, water chemistry, and equipment condition makes purely empirical models difficult to interpret and where a single failure can cascade through interconnected treatment stages.

Despite this promise, the adoption of predictive, model-based approaches in the water sector has lagged behind other industries. Barriers include fragmented and inconsistent data, limited in-house data-science capacity, integration difficulties between operational technology and information systems, and, most importantly for this study, the opacity of many machine-learning models (Carvalho et al., 2019). Because water utilities are accountable to regulators and the public, they are often reluctant to act on predictions they cannot explain. The background of this study therefore reflects both the technical maturity of digital twins and machine learning and the institutional need for a framework in which predictions are transparent, trustworthy, and embedded in lifecycle decision-making (Lockey et al., 2021).

Problem Statement

The central problem addressed by this study is that condition monitoring in U.S. drinking water treatment infrastructure remains largely reactive, schedule-based, and opaque, so that predictable asset failures may fail to be detected and prevented when digital twin modeling, XAI interpretability, sensor data validation, predictive detection, and operator engagement are weakly implemented. Although the technical capability to model assets and predict failures has advanced substantially, its translation into trustworthy, adoptable, utility-scale practice has lagged. Predictive models, when used at all, often function as black boxes whose reasoning operators cannot interrogate, which limits trust, complicates regulatory justification, and impedes adoption (Rudin, 2019). This problem matters because drinking water utilities operate critical infrastructure under strong accountability requirements, and because the consequences of undetected failures include service interruptions, water-quality violations, water loss, and public-health risk. A framework that produces accurate predictions but cannot explain them may still fail to be adopted, while a framework that is interpretable but poorly integrated with physical models and validated data may produce explanations that are misleading. The study therefore investigates how an integrated XAI digital twin framework influences the quality and reliability of predictive failure detection and lifecycle management, and which technical and institutional factors most strongly shape framework performance.

Objectives of the Study

The general objective of this study was to assess how an XAI Digital Twin Framework, integrating physics-informed digital twin modeling with interpretable machine learning, influences lifecycle management performance for U.S. drinking water treatment infrastructure. The specific objectives were: to examine the influence of digital twin modeling and integration on framework design quality; to assess the relationship between XAI model interpretability and lifecycle management performance; to evaluate the effect of sensor data quality and validation on lifecycle management performance; to determine the role of predictive failure detection accuracy in strengthening lifecycle management reliability; to test the effect of framework design quality on lifecycle management performance; and to examine the role of operator and institutional engagement in improving lifecycle outcomes. A further objective was to identify framework maturity levels and asset risk-control gaps across the management lifecycle.

Research Hypotheses

Six hypotheses guided the study. H1 proposed that digital twin modeling and integration has significantly influenced framework design quality. H2 proposed that XAI model interpretability has a significant positive relationship with lifecycle management performance. H3 proposed that sensor data quality and validation have significantly influenced lifecycle management performance. H4 proposed that predictive failure detection accuracy has significantly improved lifecycle management reliability.

H5 proposed that framework design quality has significantly predicted lifecycle management performance. H6 proposed that operator and institutional engagement has significantly strengthened lifecycle management performance.

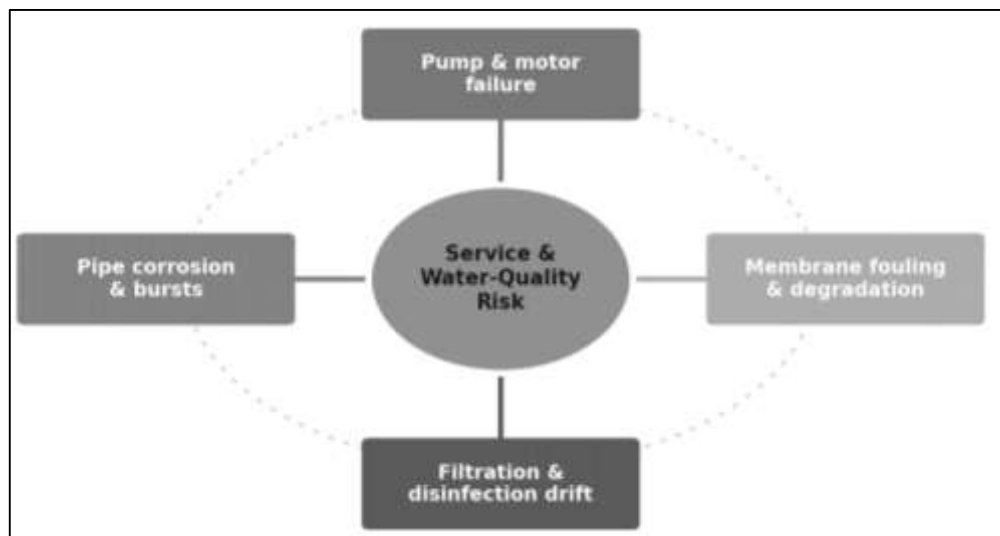
Significance of the Research

This research is significant because it addresses a practical and consequential gap between the technical maturity of digital twins and machine learning and the institutional need for transparent, trustworthy, and adoptable predictive maintenance in the water sector (Shneiderman, 2020). By treating digital twin modeling and XAI interpretability as components of a single lifecycle framework, the study offers a way to move beyond opaque, black-box prediction toward a transparent approach that utilities can adopt, justify, and defend. The findings are relevant to water utilities, control and instrumentation engineers, data scientists, asset managers, and regulators who must make high-stakes maintenance and investment decisions. The study also contributes to the broader literature on trustworthy AI for critical infrastructure by empirically examining which technical and institutional factors most strongly influence framework performance, and by developing a framework maturity index and an asset risk-control priority matrix that translate statistical results into actionable improvement priorities.

LITERATURE REVIEW

The literature on drinking water infrastructure consistently emphasizes that treatment and distribution assets face a distinctive degradation profile shaped by hydraulics, water chemistry, mechanical wear, and age. Pumps and motors degrade through bearing wear, cavitation, and electrical faults; membranes foul and lose flux; filtration media clog and channel; disinfection systems drift out of specification; and buried pipes corrode, crack, and burst (Rajani & Kleiner, 2001). Studies of pipe failure show that breakage rates rise with age, material type, soil conditions, and operating pressure, and that failures are frequent enough across the national inventory to represent a persistent operational and financial burden (Folkman, 2018). Research on these assets repeatedly stresses the interconnected nature of treatment systems. A fault in one unit process can propagate to others, degrading finished-water quality or forcing shutdowns, and the condition of any single asset is difficult to infer from sparse inspection alone (Nam et al., 2019). The literature therefore supports a systems perspective in which asset condition, water quality, and hydraulic behavior are understood together, rather than as separate concerns assessed independently, which is precisely the perspective a digital twin is designed to support.

Figure 2: Interacting Failure Modes in Drinking Water Treatment Systems



Digital Twins for Water and Infrastructure Systems

Digital twin technology has become a central concept for the condition monitoring of engineered systems. A digital twin is a virtual replica of a physical asset that is continuously synchronized with it through data, enabling simulation, monitoring, and prediction (Grieves & Vickers, 2017). The maturity

of a twin is often described along a spectrum from a digital model, through a digital shadow, to a fully connected digital twin with bidirectional data flow (Kritzinger et al., 2018). In the water sector, digital twins have been applied to distribution networks, treatment processes, and urban drainage, where they integrate hydraulic and water-quality models with real-time sensor data to support operation and planning (Conejos Fuertes et al., 2020).

The literature is candid about the requirements and limitations of digital twins. Their fidelity depends on the accuracy of the underlying physical models, the density and reliability of the sensors that update them, and the computational infrastructure that supports real-time synchronization (Rasheed et al., 2020). Recent water-sector applications show both the promise of twins for anticipating problems and the practical challenges of data integration, calibration, and maintenance of the models over time (Pesantez et al., 2022). Real-time control frameworks for water systems further demonstrate that the value of a twin is realized only when its outputs are connected to decision-making and action (Bartos & Kerkez, 2021). The reliability of a twin-based assessment therefore depends heavily on modeling choices, data quality, and integration, all of which the present study treats as elements of modeling and integration quality.

Figure 3: Digital Twin Maturity and Data-Synchronization Workflow



Explainable AI and Model Interpretability

A recurring theme in the machine-learning literature is the tension between predictive accuracy and interpretability. Complex models such as deep neural networks and gradient-boosted ensembles often achieve high accuracy but function as black boxes, while simpler models are more transparent but sometimes less accurate (Guidotti et al., 2018). Explainable AI methods address this tension by providing human-interpretable accounts of model behavior. Post-hoc methods such as local surrogate explanations (Ribeiro et al., 2016) and additive feature-attribution methods (Lundberg & Lee, 2017) estimate the contribution of each input to a prediction, while inherently interpretable models embed transparency in their structure (Rudin, 2019).

Additive feature-attribution methods make this notion of explanation precise. For a prediction on a given asset, an explanation model g assigns to each input feature a contribution ϕ_k such that the contributions sum to the model output, as shown in Equation (2). The Shapley value ϕ_k averages a feature's marginal contribution over all possible orderings of the feature set F , which gives the attribution its desirable consistency and local-accuracy properties:

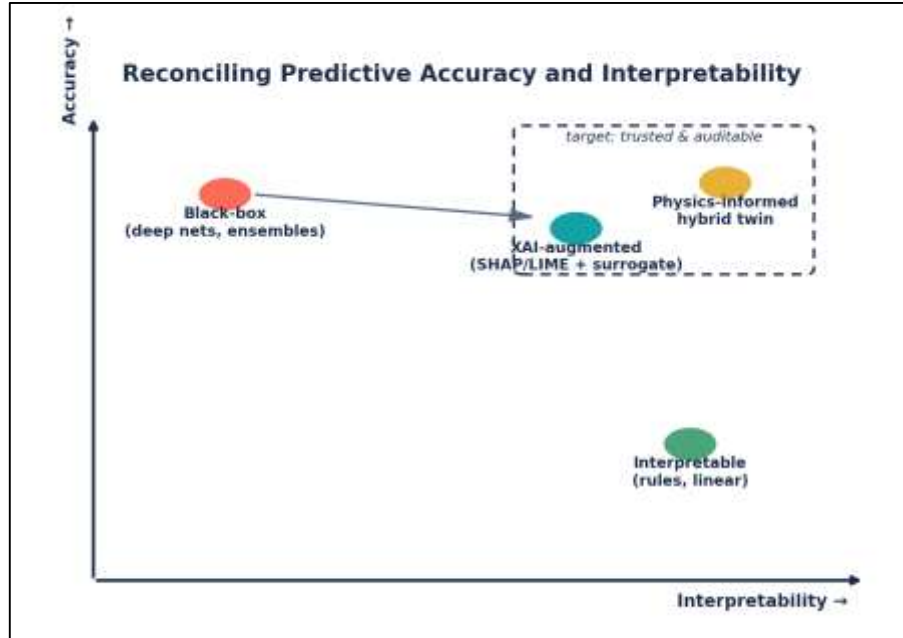
$$\hat{g}(z) = \phi_0 + \sum_{k \in F} \phi_k z_k \quad \phi_k = \sum_{S \subseteq F \setminus \{k\}} \frac{|S|! (|F| - |S| - 1)!}{(|F| - 1)!} [\hat{g}(z \cup \{k\}) - \hat{g}(z)] \quad (2)$$

In Equation (2), z -prime denotes the simplified (presence/absence) representation of the features, ϕ_0 is the base value, and the combinatorial weight ensures that each subset S of features is counted fairly. For a water utility, such attributions answer the operator's central question — which measured conditions drove this failure prediction — and thereby convert an opaque score into an auditable, defensible explanation.

The literature also emphasizes that interpretability is not merely a technical property but a socio-technical requirement, particularly in high-stakes domains. Explanations must be meaningful to the people who use them, and their adequacy depends on the audience, the task, and the consequences of

error (Barredo Arrieta et al., 2020). For critical infrastructure such as drinking water, the ability to explain why a model predicts a failure is essential for operator trust, regulatory justification, and accountability (Adadi & Berrada, 2018). The present study therefore treats XAI interpretability as a distinct construct whose inclusion transforms accurate prediction into trustworthy, adoptable decision support.

Figure 4: Reconciling Predictive Accuracy and Interpretability in the Framework



Predictive Maintenance and Physics-Informed Learning

The predictive maintenance literature describes a progression from reactive and preventive maintenance toward condition-based and predictive strategies that use data to anticipate failures (Lee et al., 2014). Machine-learning approaches to prognostics and health management have been applied across many industries to estimate remaining useful life and detect incipient faults (Zhang et al., 2019), and systematic reviews document their rapid uptake within Industry 4.0 maintenance programs (Zonta et al., 2020). Reviews of the field emphasize that the success of predictive maintenance depends not only on model accuracy but also on data quality, integration with maintenance workflows, and organizational readiness (Carvalho et al., 2019).

Physics-informed and hybrid modeling has emerged as a way to combine the strengths of physical models and data-driven learning. By embedding governing equations or physical constraints into machine-learning models, physics-informed approaches can improve generalization, reduce data requirements, and enhance interpretability (Raissi et al., 2019). Reviews of physics-informed machine learning describe its rapid growth and its particular suitability for engineered systems where physical laws are known but data are limited (Karniadakis et al., 2021). Surveys of physics-guided learning further show how domain knowledge can be integrated with data-driven models to produce predictions that are both accurate and consistent with physical understanding (Willard et al., 2022), which is directly relevant to the physics-informed digital twins that underpin the present framework. Physics-informed learning operationalizes this idea through a composite loss function that penalizes both prediction error and violation of governing physical laws. As expressed in Equation (3), the total loss combines a data-fitting term over N labeled observations with a physics-residual term evaluated at M collocation points, balanced by a weight lambda:

$$\mathcal{L}(\theta) = \underbrace{\frac{1}{N} \sum_{i=1}^N \|\hat{y}_i - y_i\|^2}_{\text{data loss}} + \lambda \underbrace{\frac{1}{M} \sum_{j=1}^M \|\mathcal{R}(\hat{y}(t_j))\|^2}_{\text{physics residual}} \quad (3)$$

The residual operator $\mathcal{R}$ encodes the governing equations of the asset, such as mass and energy balances across a treatment train or hydraulic continuity in a distribution network, so that a low residual indicates predictions consistent with physical reality. By constraining the model to respect these laws, physics-informed twins generalize better to unseen operating conditions and require less training data (Karniadakis et al., 2021), both of which are decisive advantages in water utilities where labeled failure data are scarce.

Degradation modeling links the twin's state to the timing of failure. If cumulative damage D accumulates as a function of an operating-stress signal $s(\tau)$, the remaining useful life at time t is the additional time until damage reaches a failure threshold, as defined in Equation (4):

$$D(t) = D_0 \exp\left(\int_0^t \beta s(\tau) d\tau\right), \quad \text{RUL}_t = \inf\{\Delta \geq 0 : D(t + \Delta) \geq D_{\text{th}}\} \quad (4)$$

From this degradation trajectory, the probability of failure over a prediction horizon follows from the hazard function $h(\tau)$. Assuming a Weibull hazard with shape γ and scale η , the probability that the asset fails within the interval from t to t plus Δ is given in Equation (5):

$$P_f(t, t + \Delta) = 1 - \exp\left(-\int_t^{t+\Delta} h(\tau) d\tau\right), \quad h(\tau) = \frac{\gamma}{\eta} \left(\frac{\tau}{\eta}\right)^{\gamma-1} \quad (5)$$

Equations (4) and (5) connect the physics-informed twin to actionable risk: the shape parameter γ captures whether failure risk is increasing, constant, or decreasing with age, while the scale parameter η sets the characteristic life of the component. Together they allow the framework to translate a continuously updated condition estimate into a calibrated probability of failure that operators can act on.

Figure 5: Failure Risk as the Intersection of Condition, Criticality, and Consequence



Theoretical Framework: A Dynamic Cognitive-Digital-Institutional (CDI) Loop

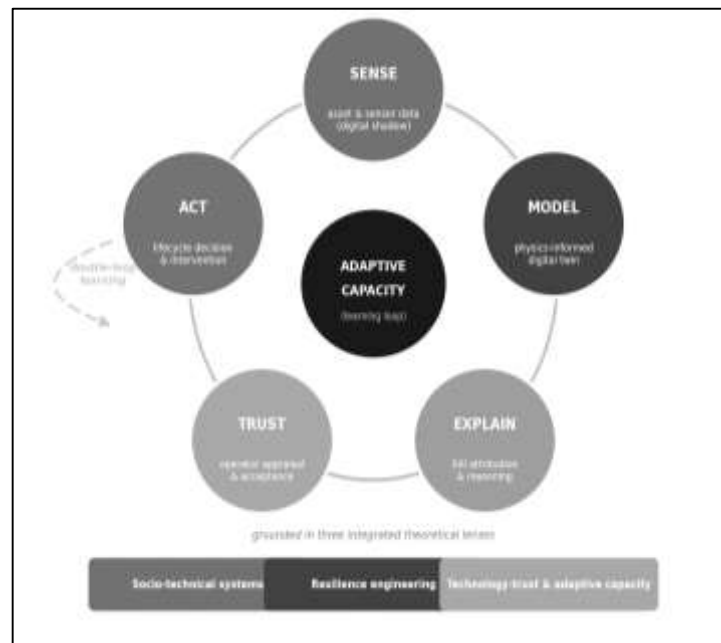
Rather than treating the framework as a static arrangement of technical and social components, this study advances a dynamic theoretical model in which predictive asset management is understood as a continuous, adaptive loop. The proposed Cognitive-Digital-Institutional (CDI) loop integrates three complementary theoretical lenses: socio-technical systems theory, which holds that technical and human subsystems must be jointly optimized rather than treated separately (Trist, 1981); resilience engineering, which frames reliable performance as a system's ongoing capacity to anticipate, monitor, respond to, and learn from disturbance (Hollnagel et al., 2006); and technology-trust and adaptive-capacity theory, which explains whether and how decision-support outputs are accepted and used by the people who rely on them (Davis, 1989). The novelty of the CDI loop is that it treats these lenses not

as parallel justifications but as interacting forces that circulate through the framework in a repeating cycle.

The loop is organized around five interacting phases that continuously feed one another. In the sense phase, physical assets and their sensors generate condition data that populate a digital shadow of the system (Kanti & Rony, 2022; Sadia et al., 2022). In the model phase, a physics-informed digital twin assimilates these data to represent asset behavior and forecast degradation. In the explain phase, XAI methods translate model outputs into human-interpretable attributions and reasoning (Barredo Arrieta et al., 2020). In the trust phase, operators appraise these explanations against their own experience, and their acceptance determines whether the prediction is treated as credible (Lockey et al., 2021). In the act phase, trusted predictions are converted into lifecycle decisions and interventions that change the physical asset, which in turn alters the data sensed in the next cycle (Tohidul, 2023; Badrul & Mominul, 2023). Because each phase both depends on and reshapes the others, the framework is inherently dynamic rather than linear, and its performance emerges from the quality of circulation around the whole loop rather than from any single stage (Hossan & Adar, 2023; Risha & Khalid, 2023).

At the center of the loop sits adaptive capacity, the system's ability to learn and reconfigure over time. Beyond the primary cycle, the model incorporates a double-loop learning mechanism in which recurring prediction errors, operator disagreements, or changing conditions feed back not only into routine action but into the retraining of models, the recalibration of twins, and the revision of institutional routines (Weick & Sutcliffe, 2007). This double-loop feature distinguishes the CDI framework from conventional predictive-maintenance pipelines, which typically treat the model as fixed once deployed (Hossan, 2024; Shima Ali et al., 2024). By making learning explicit and continuous, the framework accommodates concept drift in sensor data, evolving failure modes in aging infrastructure, and shifts in operator trust as the system demonstrates reliability over time (Shneiderman, 2020).

Figure 6: The Dynamic Cognitive-Digital-Institutional (CDI) Loop



The dynamic CDI loop therefore justifies the study's treatment of both technical constructs, such as digital twin modeling and XAI interpretability, and human-institutional constructs, such as validation and operator engagement, as mutually dependent determinants of lifecycle management performance rather than as separate inputs (Arif Uz et al., 2025; Zaman, 2024).

This double-loop learning can be formalized as a parameter-update rule, given in Equation (10). At learning step n , the framework updates its model parameters by a gradient-descent correction on the loss over the current data, augmented by a trust-driven adjustment term that encodes operator

feedback:

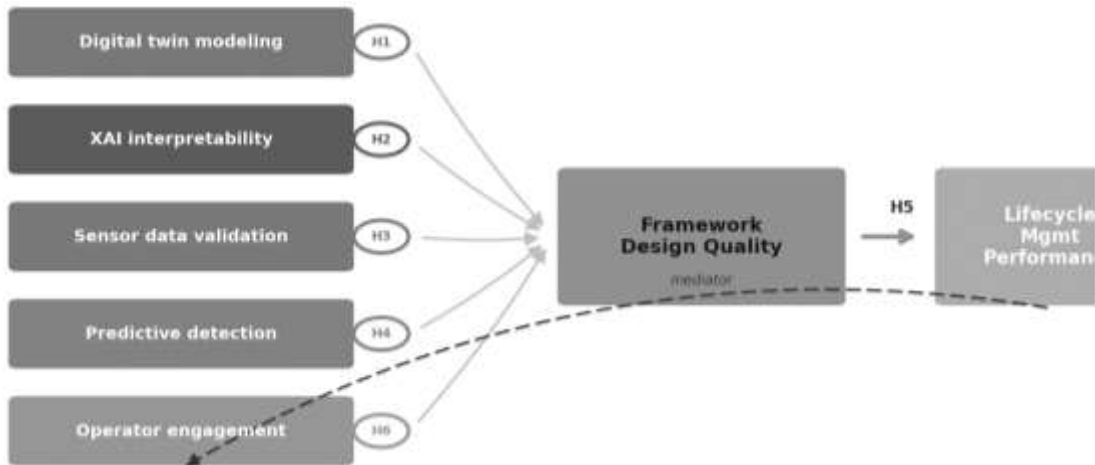
$$\theta_{n+1} = \theta_n - \eta \nabla_{\theta} \mathcal{L}(\theta_n; \mathcal{D}_n) + \kappa \Delta_{\text{trust}}^{(n)} \quad (10)$$

In Equation (10), eta is the learning rate applied to the data-driven gradient, while the second term scales an operator-trust signal by a coefficient kappa, allowing systematic operator disagreement to reshape the model rather than being discarded. This coupling of statistical learning with human feedback is what makes the CDI loop adaptive: the framework improves not only from new data but from the evolving judgment of the people who use it.

Conceptual Framework

Drawing on the reviewed literature and theoretical perspectives, the conceptual framework of this study positions lifecycle management performance as the dependent variable, influenced by six independent and mediating constructs. Digital twin modeling and integration and XAI model interpretability represent the core technical capabilities of the framework. Sensor data quality and validation, together with predictive failure detection accuracy, represent the methodological rigor that converts raw data into trustworthy, actionable predictions (Hill & Minsker, 2009). Operator and institutional engagement represents the socio-technical context that determines whether outputs are trusted and acted upon, consistent with the integrated asset-management processes that municipal utilities increasingly adopt (Halfawy, 2008) and with broader efforts toward smart, sustainable water networks (Ramos et al., 2023). Framework design quality functions as an integrating construct that captures how coherently these elements are combined, and it is expected to be the strongest direct predictor of lifecycle management performance (Mohammad Shoriful Hossan, 2025; Mukut Kanti, 2025). The framework proposes that stronger implementation of each construct, circulating together through the dynamic CDI loop, produces more reliable, transparent, and actionable predictive asset management for drinking water utilities.

Figure 7: Predictive Conceptual Framework for Lifecycle Management Performance



METHODS

This study adopted a quantitative, cross-sectional, case-based research design to examine how an XAI Digital Twin Framework, integrating physics-informed digital twin modeling with interpretable machine learning, influences lifecycle management performance in U.S. drinking water treatment infrastructure. The quantitative approach was selected because the study measured relationships among key variables using numerical data collected through a structured questionnaire. The cross-sectional design was appropriate because data were gathered from respondents at a single point in time, while the case-based approach enabled the research to focus on drinking water utilities where digital twins, machine learning, and predictive maintenance have been applied or considered relevant for asset management.

The case context focused on drinking water treatment and distribution environments such as surface-water and groundwater treatment plants, membrane and filtration facilities, pumping stations,

disinfection systems, and distribution networks. These settings were selected because they depend heavily on condition monitoring for the reliability of pumps, membranes, filters, disinfection equipment, sensors, and pipes. The study considered these environments suitable because an XAI digital twin framework is highly relevant to their predictive maintenance, lifecycle planning, and regulatory-compliance needs. The population of the study consisted of professionals involved in water treatment operations, control and instrumentation engineering, data science, asset management, maintenance planning, and utility governance. The unit of analysis was the individual professional respondent who possessed knowledge or experience related to digital twins, machine learning, predictive maintenance, sensor systems, or asset management in the water sector. The study used a purposive sampling strategy because respondents needed specific technical or institutional knowledge to provide valid responses. Where possible, the sample included water treatment operators, control and instrumentation engineers, data scientists and machine-learning engineers, asset-management and maintenance planners, and utility managers and regulatory staff.

The data-collection procedure involved a structured questionnaire distributed to selected respondents electronically or physically. Before data collection, respondents were informed about the academic purpose of the study, voluntary participation, confidentiality, and anonymity. The collected responses were reviewed for completeness, consistency, and suitability before analysis, and incomplete or invalid responses were excluded to maintain data quality. The instrument design was based on a five-point Likert scale ranging from 1 = Strongly Disagree to 5 = Strongly Agree. The questionnaire included sections on demographic profile, digital twin modeling and integration, XAI model interpretability, sensor data quality and validation, predictive failure detection accuracy, operator and institutional engagement, framework design quality, and lifecycle management performance.

A pilot test was conducted before the main survey to assess the clarity, relevance, wording, and technical suitability of the questionnaire items. Feedback from the pilot test was used to improve ambiguous or repetitive items. Validity and reliability were ensured through expert review and statistical testing. Content validity was supported by aligning questionnaire items with the study objectives and predictive-maintenance concepts, while reliability was tested using Cronbach's alpha, with a value of 0.70 or above considered acceptable.

Internal consistency was quantified using Cronbach's alpha, defined in Equation (6), where K is the number of items in a construct, the numerator sums the individual item variances, and the denominator is the variance of the summed scale:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K \sigma_i^2}{\sigma_T^2} \right) \quad (6)$$

Values approaching unity indicate that the items measure a common underlying construct, and the 0.70 threshold reflects the conventional lower bound for acceptable reliability. For data analysis, statistical software was used for coding, descriptive statistics, reliability analysis, correlation analysis, and regression modeling. A spreadsheet application was used for initial data cleaning, tabulation, and organization of responses, and reference-management software was used for citation organization and APA formatting. Two model specifications underpinned the analysis. First, framework design quality was operationalized as a weighted composite of its constituent construct means, as shown in Equation (7), where each construct mean is weighted by a non-negative weight summing to one:

$$EDQ = \sum_C w_C \bar{y}_C \quad \sum_C w_C = 1 \quad w_C \geq 0 \quad (7)$$

Second, the influence of the six predictors on lifecycle management performance was estimated with the multiple linear regression model in Equation (8), where LMP denotes lifecycle management performance, each X_k is a predictor construct, and the error term is assumed normally distributed with zero mean and constant variance:

$$LMP = \beta_0 + \sum_{k=1}^6 \beta_k X_k + \epsilon \quad \epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2) \quad (8)$$

The standardized coefficients β_k reported in the regression results express the relative predictive strength of each construct, and the coefficient of determination R -squared quantifies the proportion of

variance in lifecycle management performance jointly explained by the six predictors.

FINDINGS

Response Rate

The response-rate results in Table 1 showed that the study achieved a strong and acceptable level of participation from professionals involved in drinking water asset management. Out of 155 questionnaires distributed, 145 responses were returned, representing a return rate of 93.5%. After screening for completeness, consistency, and suitability for statistical analysis, 7 responses were removed because they contained missing values, incomplete Likert-scale answers, or inconsistent response patterns. As a result, 138 valid responses were retained for final analysis, producing an effective valid response rate of 89.0%.

Table 1: Response Rate of the Survey

Survey Status	Frequency	Percentage
Questionnaires distributed	155	100.0%
Questionnaires returned	145	93.5%
Invalid/incomplete responses	7	4.5%
Valid responses used for analysis	138	89.0%

This response rate was considered suitable for a quantitative, cross-sectional, case-based study because it provided enough data to conduct descriptive statistics, reliability analysis, correlation analysis, and regression modeling. The response rate also strengthened the credibility of the findings because the respondents represented relevant professional groups connected to water treatment operations, control and instrumentation, data science, asset management, maintenance planning, and utility governance. From a socio-technical perspective, the strong response rate suggested that the selected professional population showed meaningful engagement with condition-monitoring issues and provided sufficient institutional insight into technical capability, model interpretability, validation practices, and operator engagement. The valid sample of 138 respondents therefore provided a dependable foundation for examining the study objectives and hypotheses, allowing the research to test whether digital twin modeling, XAI interpretability, sensor validation, predictive detection, operator engagement, and framework design quality significantly contributed to lifecycle management performance.

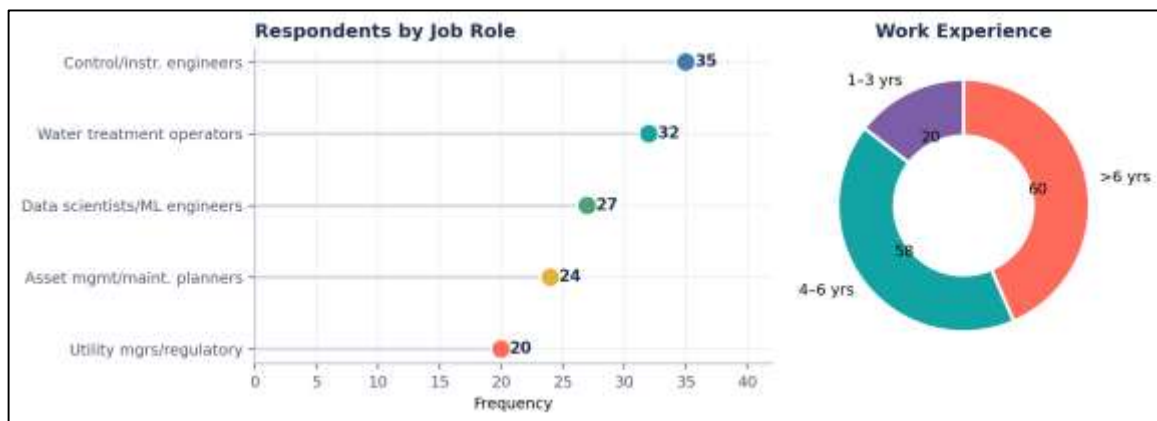
Demographic and Professional Profile of Respondents

Table 2 presented the demographic and professional characteristics of the 138 respondents included in the analysis. The results indicated that the sample was technically relevant to the research topic because the largest group of respondents consisted of control and instrumentation engineers, accounting for 25.4% of the sample. This was followed by water treatment operators at 23.2%, data scientists and machine-learning engineers at 19.6%, asset-management and maintenance planners at 17.4%, and utility managers and regulatory staff at 14.5%. This distribution was useful because an XAI digital twin framework depends on multiple professional roles rather than one technical group alone. Control and instrumentation engineers are typically involved in sensor systems, control logic, and equipment monitoring. Data scientists and machine-learning engineers contribute expertise in model development, interpretability, and validation. Water treatment operators contribute practical knowledge of asset behavior, alarms, and day-to-day condition. Asset managers and planners contribute knowledge of lifecycle planning, prioritization, and investment. Utility managers and regulatory staff contribute governance, accountability, and compliance perspectives. The experience profile also strengthened the study because 43.5% of respondents had more than six years of experience, while 42.0% had four to six years of experience. In addition, 65.9% of respondents reported direct involvement in condition-monitoring or asset-management activities, while 34.1% had indirect involvement.

Table 2: Demographic and Professional Profile of Respondents

Profile Variable	Category	Frequency	Percentage
Job role	Control/instrumentation engineers	35	25.4%
Job role	Water treatment operators	32	23.2%
Job role	Data scientists/ML engineers	27	19.6%
Job role	Asset-management/ maintenance planners	24	17.4%
Job role	Utility managers/regulatory staff	20	14.5%
Work experience	1-3 years	20	14.5%
Work experience	4-6 years	58	42.0%
Work experience	Above 6 years	60	43.5%
Involvement	Direct involvement	91	65.9%
Involvement	Indirect involvement	47	34.1%

This profile supported the validity of the findings because the study depended on respondents who understood digital twin modeling, model interpretability, sensor validation, predictive detection, and lifecycle management. From a socio-technical perspective, the sample reflected a cross-functional system in which reliable asset management depends on communication, expertise, technical awareness, and coordination among different professional groups. The respondent profile was therefore appropriate for addressing the study objectives and testing the hypotheses.

Figure 12: Respondent Profile by Job Role and Work Experience

Descriptive Statistics of Study Variables

Table 3: Descriptive Statistics of Study Variables Based on Five-Point Likert Scale

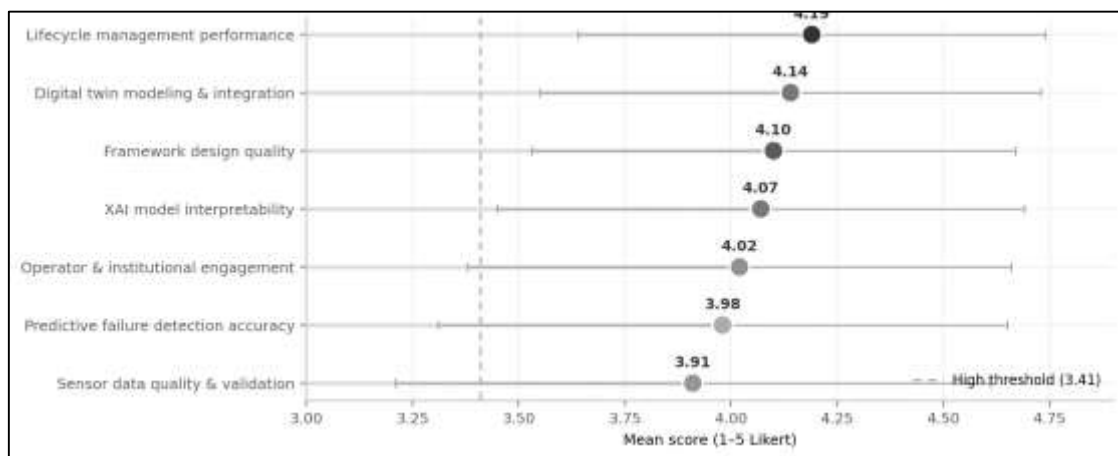
Study Variable	Mean	SD	Interpretation	Rank
Lifecycle management performance	4.19	0.55	High	1
Digital twin modeling and integration	4.14	0.59	High	2
Framework design quality	4.10	0.57	High	3
XAI model interpretability	4.07	0.62	High	4
Operator and institutional engagement	4.02	0.64	High	5
Predictive failure detection accuracy	3.98	0.67	High	6
Sensor data quality and validation	3.91	0.70	High	7

Likert interpretation guide: 1.00–1.80 = Very Low, 1.81–2.60 = Low, 2.61–3.40 = Moderate, 3.41–4.20 = High, 4.21–5.00 = Very High.

Table 3 showed the descriptive statistics for the major study variables measured through the five-point Likert scale. All variables recorded mean scores above the neutral midpoint of 3.00 and within the high interpretation range of 3.41–4.20, suggesting that respondents generally agreed that XAI digital twin practices, digital twin modeling, model interpretability, sensor validation, predictive detection, operator engagement, and lifecycle management were positively present in the selected water-utility

context. Lifecycle management performance recorded the highest mean of 4.19 with a standard deviation of 0.55, indicating that respondents perceived early failure detection, condition characterization, intervention prioritization, and repeatable maintenance decisions as relatively strong. Digital twin modeling and integration recorded a mean of 4.14, showing that virtual modeling, data assimilation, and simulation were perceived as well implemented. Framework design quality followed with a mean of 4.10, suggesting agreement that the coherent combination of technical and institutional elements supported management goals. XAI model interpretability recorded 4.07, operator and institutional engagement recorded 4.02, and predictive failure detection accuracy recorded 3.98. Sensor data quality and validation recorded the lowest mean of 3.91. Although all values remained in the high range, the lower ranking of validation and predictive accuracy suggested that these data-facing and methodological activities may require more focused improvement. These findings aligned with the dynamic CDI loop because the strongest lifecycle performance appeared where the sense, model, explain, trust, and act phases circulated together rather than in isolation.

Figure 13: Mean Scores of Study Variables with Standard Deviation



Reliability Analysis

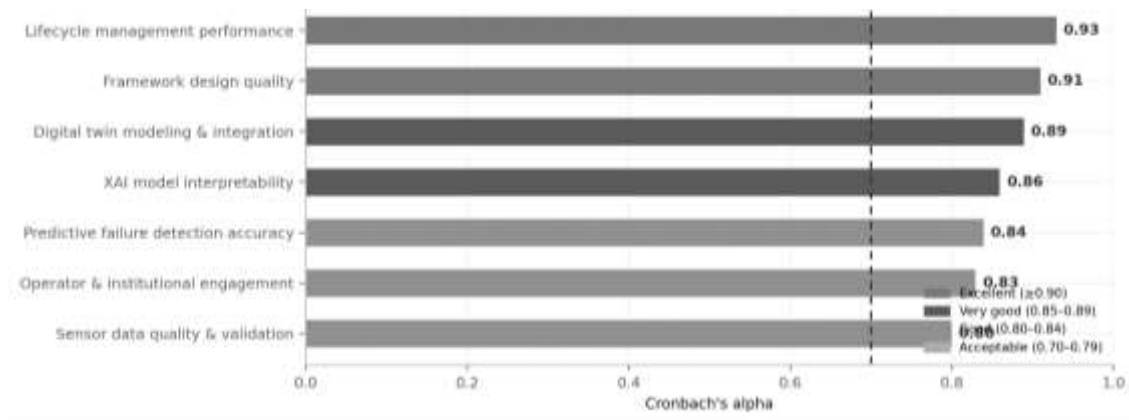
Table 4: Reliability Analysis of Study Constructs

Construct	Items	Cronbach's Alpha	Decision
Digital twin modeling and integration	6	0.89	Reliable
XAI model interpretability	6	0.86	Reliable
Sensor data quality and validation	5	0.80	Reliable
Predictive failure detection accuracy	5	0.84	Reliable
Operator and institutional engagement	5	0.83	Reliable
Framework design quality	6	0.91	Reliable
Lifecycle management performance	6	0.93	Reliable

Table 4 presented the reliability results for the major questionnaire constructs. Cronbach's alpha was used to determine whether the items measuring each construct were internally consistent. The accepted threshold was 0.70, and all constructs exceeded this minimum. Lifecycle management performance recorded the highest alpha of 0.93, indicating excellent internal consistency among items measuring early detection, condition characterization, prioritization, and repeatable maintenance decisions. Framework design quality recorded 0.91, confirming that items measuring the coherent integration of technical and institutional elements formed a reliable construct. Digital twin modeling and integration recorded 0.89, XAI model interpretability recorded 0.86, and predictive failure detection accuracy recorded 0.84. Operator and institutional engagement recorded 0.83, confirming that items on communication, trust, and institutional coordination reliably measured the socio-technical dimension. Sensor data quality and validation recorded the lowest alpha of 0.80, which remained well above the accepted threshold and was therefore considered reliable. These reliability results strengthened the

trustworthiness of the study by showing that the questionnaire measured the intended constructs consistently, and they supported the socio-technical view that lifecycle performance is measurable as a combination of technical and institutional practices.

Figure 14: Internal Consistency Reliability (Cronbach's Alpha) of Study Constructs



Correlation Analysis

Table 5: Pearson Correlation Analysis among Major Study Variables

Relationship Tested	r	Significance	Interpretation	Link
Digital twin modeling and framework design quality	0.74	$p < 0.01$	Strong positive	H1
XAI interpretability and lifecycle management performance	0.68	$p < 0.01$	Strong positive	H2
Data quality/validation and lifecycle performance	0.71	$p < 0.01$	Strong positive	H3
Predictive detection accuracy and reliability	0.64	$p < 0.01$	Moderate-to-strong	H4
Operator engagement and lifecycle performance	0.66	$p < 0.01$	Strong positive	H6
Framework design quality and lifecycle performance	0.77	$p < 0.01$	Strong positive	H5

Table 5 presented the Pearson correlation results used to examine the strength and direction of relationships among the major study variables. All tested relationships were positive and statistically significant at $p < 0.01$, indicating that stronger implementation of technical and institutional practices was associated with stronger lifecycle management performance. The relationship between digital twin modeling and integration and framework design quality recorded a strong positive correlation of $r = 0.74$, supporting H1 by showing that stronger modeling capability accompanied stronger framework design. XAI model interpretability correlated with lifecycle performance at $r = 0.68$, supporting H2. Sensor data quality and validation correlated with lifecycle performance at $r = 0.71$, supporting H3 and indicating that validation and error control contributed meaningfully to trustworthy outputs. Predictive failure detection accuracy recorded $r = 0.64$ with lifecycle reliability, supporting H4. Operator and institutional engagement recorded $r = 0.66$ with lifecycle performance, supporting the socio-technical argument that reliable management depends on institutional and human behavior as well as technical systems. Framework design quality showed the strongest relationship with lifecycle performance at $r = 0.77$, supporting H5 and confirming that coherent design was highly relevant to outcomes. These findings aligned with the dynamic CDI loop because lifecycle performance improved when technical modeling, explanation, trust, and action circulated together as an integrated, adaptive cycle.

Figure 15: Correlation Strength among Major Study Variables



Regression Analysis

Table 6: Multiple Regression Results Predicting Lifecycle Management Performance

Predictor Variable	Std. Beta (β)	t-value	p-value	Decision
Digital twin modeling and integration	0.26	3.58	0.001	Significant
XAI model interpretability	0.23	3.19	0.002	Significant
Sensor data quality and validation	0.19	2.74	0.007	Significant
Predictive failure detection accuracy	0.16	2.46	0.015	Significant
Operator and institutional engagement	0.18	2.71	0.008	Significant
Framework design quality	0.33	4.61	<0.001	Significant

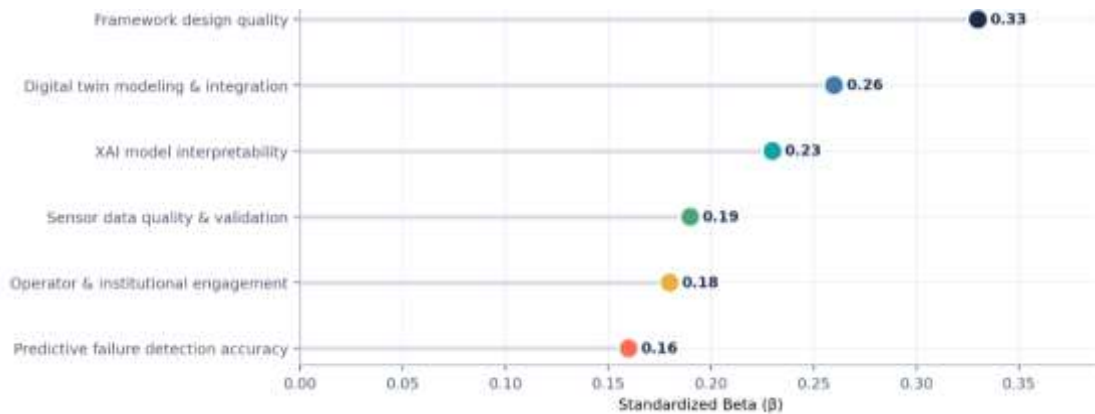
Table 7: Model Summary for Regression Analysis

Model Indicator	Value
R	0.838
R ²	0.702
Adjusted R ²	0.689
F-statistic	51.36
df	6, 131
Significance	p < 0.001

Tables 6 and 7 presented the multiple regression results used to determine the predictive effect of the independent variables on lifecycle management performance. The regression model was statistically significant, $F(6, 131) = 51.36$, $p < 0.001$, indicating that the selected predictors collectively explained lifecycle performance in a meaningful way. The model recorded an R value of 0.838, showing a strong overall relationship between the predictors and the dependent variable. The R² value of 0.702 indicated that 70.2% of the variance in lifecycle management performance was explained by digital twin modeling, XAI interpretability, data quality and validation, predictive detection, operator engagement, and framework design quality. The adjusted R² value of 0.689 confirmed that the model remained strong after accounting for the number of predictors. Among the predictors, framework design quality showed the strongest effect, with $\beta = 0.33$ and $p < 0.001$, indicating that coherent framework design was the most important predictor of lifecycle outcomes. Digital twin modeling and integration showed

a significant positive effect, with $\beta = 0.26$ and $p = 0.001$, confirming that strong modeling improved lifecycle performance. XAI model interpretability recorded $\beta = 0.23$ and $p = 0.002$, while sensor data quality and validation recorded $\beta = 0.19$ and $p = 0.007$. Operator and institutional engagement recorded $\beta = 0.18$ and $p = 0.008$, and predictive failure detection accuracy recorded $\beta = 0.16$ and $p = 0.015$, a smaller but still significant contribution. These results supported the dynamic CDI loop because lifecycle performance was explained jointly by the digital, cognitive, and institutional phases of the cycle, and they directly supported the study objectives by proving that the selected variables significantly predicted lifecycle management performance.

Figure 16: Predictive Strength of Factors on Lifecycle Management Performance (Standardized Beta)



Hypotheses Testing

Table 8: Summary of Hypotheses Testing

Hypothesis	Statement	Statistical Evidence	Decision
H1	Digital twin modeling has significantly influenced framework design quality.	$r = 0.74, \beta = 0.26, p = 0.001$	Supported
H2	XAI interpretability has a significant positive relationship with lifecycle performance.	$r = 0.68, \beta = 0.23, p = 0.002$	Supported
H3	Sensor data quality and validation have significantly influenced lifecycle performance.	$r = 0.71, \beta = 0.19, p = 0.007$	Supported
H4	Predictive detection accuracy has significantly improved lifecycle reliability.	$r = 0.64, \beta = 0.16, p = 0.015$	Supported
H5	Framework design quality has significantly predicted lifecycle performance.	$r = 0.77, \beta = 0.33, p < 0.001$	Supported
H6	Operator engagement has significantly strengthened lifecycle performance.	$r = 0.66, \beta = 0.18, p = 0.008$	Supported

Table 8 summarized the results of the hypotheses testing and showed that all six hypotheses were supported by the statistical findings. H1 was supported because digital twin modeling and integration showed a strong positive relationship with framework design quality and produced a significant regression effect, confirming that stronger modeling capability accompanied stronger framework design. H2 was supported because XAI model interpretability showed a significant positive relationship with lifecycle performance, meaning that more interpretable predictions were associated with stronger outcomes. H3 was supported because sensor data quality and validation significantly influenced lifecycle performance, confirming that error control, sensor validation, and data cleaning improved trust in results. H4 was supported because predictive failure detection accuracy showed a significant positive effect on lifecycle reliability, suggesting that accurate detection of incipient failures strengthened maintenance readiness. H5 was supported because framework design quality was the strongest predictor of lifecycle performance, with $\beta = 0.33$ and $p < 0.001$, directly proving the main objective of the study. H6 was supported because operator and institutional engagement showed a significant positive effect, linking the findings to the trust phase of the CDI loop, which holds that reliable performance depends on operator appraisal and acceptance as much as on technical modeling.

Overall, the hypotheses testing demonstrated that the study model was statistically supported and that both technical and institutional variables contributed to lifecycle management performance in drinking water utilities.

Framework Compliance Maturity Index

Table 9: XAI Digital Twin Framework Maturity Index

Framework Lifecycle Stage	Mean	SD	Maturity Level	Priority
Asset scoping and data acquisition	4.21	0.53	Very high maturity	Maintain and strengthen
Digital twin modeling and calibration	4.15	0.56	High maturity	Maintain
Predictive model development	4.12	0.58	High maturity	Maintain
Explainability and interpretation	3.97	0.66	High maturity	Improve documentation
Sensor validation and data quality	3.90	0.71	High maturity	Needs attention
Operator engagement and trust	3.87	0.73	High maturity	Needs attention
Lifecycle decision and audit readiness	3.95	0.67	High maturity	Improve consistency
Overall framework maturity	4.02	0.63	High maturity	Strong but improvable

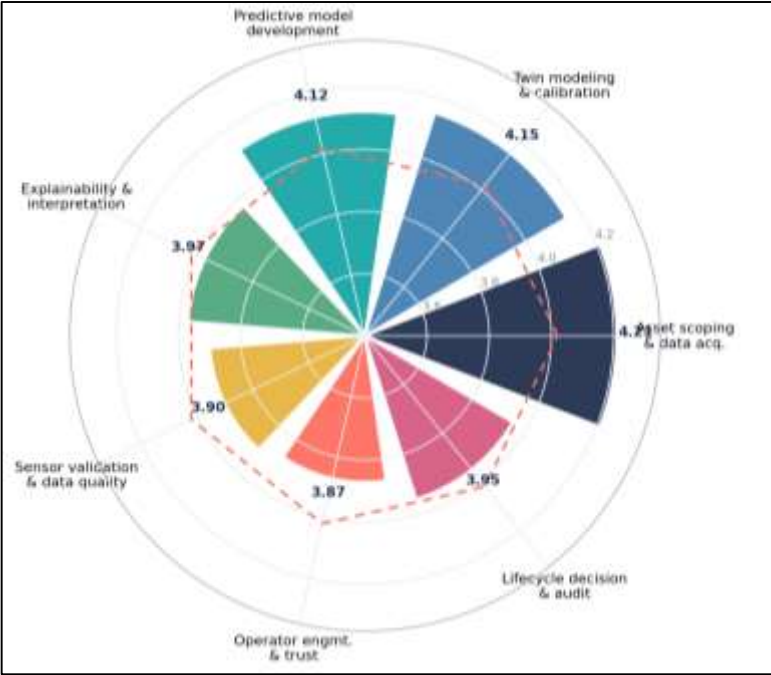
The framework maturity index (MI) and the risk-control gap score were computed as shown in Equation (9). The maturity index is the mean maturity across the S lifecycle stages, while the gap score for a risk-control area a is the difference between its importance mean and its effectiveness mean:

$$MI = \frac{1}{S} \sum_{s=1}^S m_{ss}, \quad \text{Gap}_a = \bar{I}_a - \bar{E}_a \quad (9)$$

Maturity interpretation guide: 1.00–1.80 = Very Low Maturity, 1.81–2.60 = Low Maturity, 2.61–3.40 = Moderate Maturity, 3.41–4.20 = High Maturity, 4.21–5.00 = Very High Maturity.

Table 9 presented the XAI Digital Twin Framework Maturity Index, developed as a study-specific result to assess how mature the selected water utilities were across major stages of the framework lifecycle. The overall maturity score was 4.02, falling within the high-maturity range and indicating that the case context generally demonstrated strong implementation of XAI digital twin practices. Asset scoping and data acquisition recorded the highest maturity mean of 4.21, reaching the very-high band and suggesting strong capability in defining assets, instrumenting them, and acquiring condition data. Digital twin modeling and calibration recorded 4.15, while predictive model development recorded 4.12, showing that the core technical stages were relatively strong. Explainability and interpretation recorded 3.97, indicating high maturity but room for improvement in documenting and communicating model reasoning. Lifecycle decision and audit readiness recorded 3.95, suggesting that records were available but could benefit from better standardization. Sensor validation and data quality recorded 3.90, while operator engagement and trust recorded the lowest maturity score of 3.87. Although both values remained classified as high maturity, their lower ranking showed that data-facing and human-facing activities required more attention than the core technical stages. This result aligned with the descriptive findings, where validation and predictive accuracy recorded the lowest means among the major constructs. The maturity index linked directly to the CDI loop because it showed that reliable asset management depends on continuous circulation through the sense, model, explain, trust, and act phases, supported by double-loop learning, rather than on one-time model development.

Figure 18: Framework Maturity Across Lifecycle Stages



Asset Risk-Control Priority Matrix

Table 10 presented the Asset Risk-Control Priority Matrix, developed to compare the perceived importance and current implementation effectiveness of major risk-control areas within the framework. The gap score was calculated by subtracting the effectiveness mean from the importance mean, with a higher gap score indicating a higher improvement priority. The results showed that sensor validation and calibration recorded the largest gap score of 1.13, followed closely by operator trust and explanation quality with a gap score of 1.09. Both areas were therefore classified as very high priority, meaning that respondents considered them highly important while rating their current effectiveness lower than desired. Model interpretability and transparency was also classified as very high priority, with an importance mean of 4.63, an effectiveness mean of 3.70, and a gap score of 0.93, suggesting a need to strengthen the link between accurate prediction and human-understandable explanation. Membrane and pump degradation detection and remaining-useful-life estimation both recorded gap scores in the high-priority range.

Table 10: Asset Risk-Control Priority Matrix

Asset Risk-Control Area	Importance	Effectiveness	Gap	Priority
Pump and motor failure detection	4.74	3.97	0.77	High priority
Membrane fouling and degradation	4.66	3.72	0.94	Very high priority
Sensor validation and calibration	4.60	3.47	1.13	Very high priority
Operator trust and explanation quality	4.52	3.43	1.09	Very high priority
Model interpretability and transparency	4.63	3.70	0.93	Very high priority
Remaining-useful-life estimation	4.40	3.68	0.72	High priority
Distribution asset condition integration	4.37	3.76	0.61	Medium priority
Data governance and model updating	4.31	3.61	0.70	High priority

Distribution asset condition integration recorded the lowest gap score of 0.61 and was classified as medium priority, suggesting that it was relatively stronger than other areas but still required continued monitoring. Remaining-useful-life estimation recorded 0.72 and data governance and model updating recorded 0.70, both classified as high priority. This priority matrix strengthened the findings by translating statistical results into practical improvement priorities. It also aligned with socio-technical and resilience perspectives, which emphasize attention to weak signals, human trust, and continuous improvement in high-consequence environments. The matrix showed that reliable performance required attention not only to core technical modeling but also to sensor validation, model interpretability, and operator trust, which supported the study objectives by identifying where framework improvements should be prioritized.

Figure 19: Asset Risk-Control Importance versus Effectiveness Gap Analysis

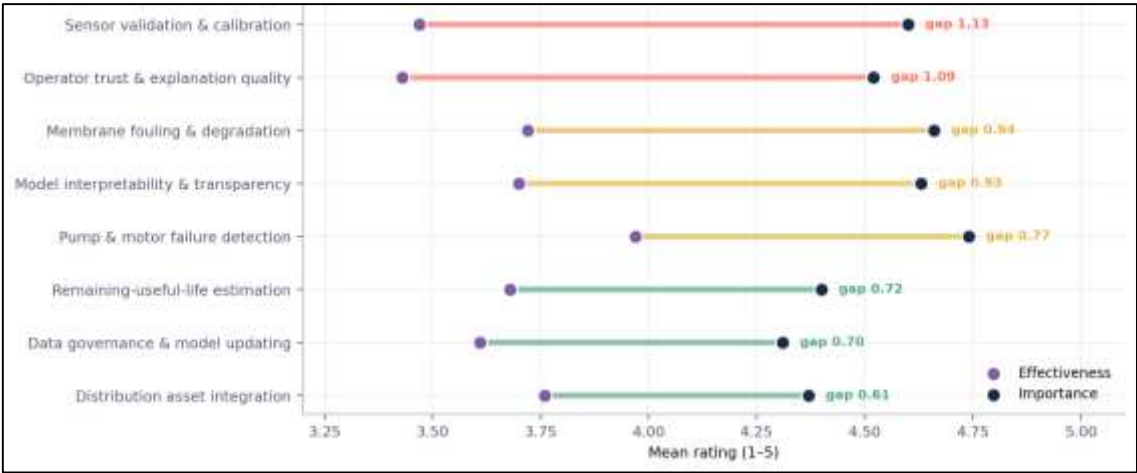


Table 11: Summary of Key Findings by Objective and Hypothesis

Study Objective	Main Result	Hypothesis	Interpretation
Examine influence of digital twin modeling on framework design quality	Twin modeling strongly correlated with design quality, $r = 0.74$	H1	Stronger modeling improved design quality
Assess relationship between XAI interpretability and lifecycle performance	Interpretability significantly related to performance, $r = 0.68$, $\beta = 0.23$	H2	Interpretable models improved outcomes
Evaluate effect of sensor data quality and validation	Validation significantly predicted performance, $\beta = 0.19$	H3	Validation strengthened trust in results
Determine role of predictive detection accuracy	Detection accuracy significantly predicted reliability, $\beta = 0.16$	H4	Accurate detection supported readiness
Test effect of framework design quality on performance	Design quality was the strongest predictor, $\beta = 0.33$	H5	Design quality was central to performance
Examine role of operator engagement	Engagement significantly predicted performance, $\beta = 0.18$	H6	Engagement supported resilient performance
Identify maturity and risk-control gaps	Overall maturity = 4.02; largest gaps in sensor validation and operator trust	Supports overall model	Data- and human-facing controls required improvement

Table 11 summarized the major findings in relation to the research objectives and hypotheses, showing that all objectives were achieved and all hypotheses supported. The first objective was achieved because digital twin modeling and integration showed a strong positive relationship with framework design quality, $r = 0.74$, confirming that stronger modeling capability improved the coherence of framework

design. The second objective was achieved because XAI model interpretability showed a significant positive relationship with lifecycle performance, $r = 0.68$ and $\beta = 0.23$, demonstrating that interpretable predictions contributed to stronger outcomes. The third objective was achieved because sensor data quality and validation significantly predicted lifecycle performance, $\beta = 0.19$, showing that error control and validation improved confidence in results. The fourth objective was achieved because predictive failure detection accuracy significantly predicted lifecycle reliability, $\beta = 0.16$, the smallest but still significant coefficient, showing that accurate detection contributed to management quality. The fifth objective was achieved because framework design quality was the strongest predictor, $\beta = 0.33$ and $p < 0.001$, confirming the central argument of the research. The sixth objective was supported because operator and institutional engagement significantly predicted performance, $\beta = 0.18$, linking the results to the trust phase of the dynamic CDI loop. The maturity index showed an overall score of 4.02, while the priority matrix identified sensor validation, operator trust, and model interpretability as the highest-priority improvement areas. Overall, the results proved that lifecycle management performance in drinking water utilities depended on the combined effect of digital twin modeling, XAI interpretability, sensor validation, predictive detection, operator engagement, and framework design quality.

FINDINGS

This chapter presents the findings of the quantitative analysis conducted to examine how an XAI Digital Twin Framework, integrating physics-informed digital twin modeling with interpretable machine learning, influences lifecycle management performance in U.S. drinking water treatment infrastructure. The analysis was based on data collected through a structured five-point Likert-scale questionnaire, where 1 represented Strongly Disagree and 5 represented Strongly Agree. A total of 155 questionnaires were distributed to professionals involved in water treatment operations, control and instrumentation, data science, asset management, maintenance planning, and utility governance, of which 138 valid responses were retained for analysis, producing a valid response rate of 89.0%.

The descriptive findings showed that all seven study variables were rated in the high range, with lifecycle management performance recording the highest mean of 4.19 and sensor data quality and validation recording the lowest mean of 3.91. This pattern indicated broad agreement that the framework's components were positively present, while also signaling that validation and data-facing activities were perceived as comparatively weaker. The reliability findings confirmed that all constructs were internally consistent, with Cronbach's alpha values ranging from 0.80 to 0.93, allowing the study to proceed confidently to inferential analysis.

The correlation findings showed that all six hypothesized relationships were positive and statistically significant at $p < 0.01$, with framework design quality showing the strongest relationship with lifecycle performance, $r = 0.77$, and digital twin modeling showing a strong relationship with framework design quality, $r = 0.74$. The regression findings showed that the six predictors together explained 70.2% of the variance in lifecycle performance, with framework design quality the strongest predictor, $\beta = 0.33$, followed by digital twin modeling, $\beta = 0.26$, and XAI interpretability, $\beta = 0.23$. The hypotheses-testing findings confirmed that all six hypotheses were supported. The maturity index found overall high maturity, 4.02, with the strongest maturity in asset scoping, twin modeling, and predictive development, and the weakest in validation, explainability, and operator engagement. The priority matrix identified sensor validation, operator trust, and model interpretability as the highest-priority improvement areas. Collectively, these findings demonstrate that an XAI digital twin framework can substantially strengthen predictive asset management in drinking water utilities, and that its reliability depends jointly on technical capability, methodological validation, and operator engagement.

DISCUSSION

The findings of this study provide strong empirical support for the proposition that an XAI digital twin framework substantially strengthens predictive failure detection and lifecycle management in U.S. drinking water treatment infrastructure, and that its reliability depends on the joint operation of technical capability, methodological rigor, and institutional engagement. The result that framework design quality was the strongest predictor of lifecycle performance, $\beta = 0.33$, is consistent with the study's conceptual argument that the value of digital twins and machine learning is realized not through any single technique but through the coherent integration of modeling, explanation,

validation, and operator engagement (Fuller et al., 2020). This finding echoes the broader literature on trustworthy AI for critical systems, which repeatedly emphasizes that raw predictive capability must be embedded in a well-designed, human-centered workflow to produce dependable outputs (Shneiderman, 2020). Interpreted through the dynamic CDI loop, framework design quality captures how smoothly information circulates across the sense, model, explain, trust, and act phases, so its dominance as a predictor reflects the loop's premise that performance emerges from the whole cycle rather than any single stage.

The strong relationship between digital twin modeling and integration and framework design quality, $r = 0.74$, and the significant regression effect of twin modeling on lifecycle performance, $\beta = 0.26$, confirm the central role of physics-informed virtual modeling in this context. These results align with prior studies showing that digital twins can represent the dynamic behavior of water systems at a fidelity that supports anticipatory decision-making (Conejos Fuertes et al., 2020). At the same time, the comparatively lower means and maturity scores for sensor validation and predictive accuracy suggest that respondents recognized the well-documented dependence of twins on data quality; a twin is only as reliable as the sensing and validation that feed it (Jones et al., 2020). The prominence of sensor validation in the priority matrix, where it recorded the largest gap score of 1.13, reinforces the interpretation that practitioners view data quality as essential yet under-implemented, a tension the framework must resolve to be trusted (Mukut Kanti, 2025; Shima Ali et al., 2025).

The significant contribution of XAI model interpretability, $\beta = 0.23$, provides direct empirical support for the argument that interpretability is not a technical luxury but a precondition for adoption in accountable, high-consequence settings (Rudin, 2019). By making model reasoning visible through feature attributions and interpretable structures, the framework enables operators and regulators to scrutinize and trust predictions (Lundberg & Lee, 2017). This finding is consistent with the socio-technical literature, which stresses that the adequacy of an explanation depends on the audience and the task, and that trust is built through transparency and demonstrated reliability (Lockey et al., 2021). However, the very-high-priority classification of operator trust and explanation quality in the priority matrix, with a gap of 1.09, suggests that the human-facing dimension of the framework remains the least mature: technical modeling capacity has advanced faster than the interpretive and institutional capacity needed to convert predictions into trusted action (Amir, 2026; Mahmudul & Sadia, 2026).

The significant effect of operator and institutional engagement, $\beta = 0.18$, provides further support for the trust phase of the dynamic CDI loop that underpins the study. This result indicates that even accurate, interpretable predictions do not achieve their full value unless operators are engaged and institutions are able to act on them (Davis, 1989). Technology-acceptance and resilience research both suggest that perceived usefulness, trust, and organizational readiness are decisive for whether decision-support tools are adopted and used (Weick & Sutcliffe, 2007). Taken together, the results describe a framework that is technically strong at its core, with high maturity in asset scoping, twin modeling, and predictive development, but comparatively weaker at its data-facing and human-facing margins, where validation, interpretability, and operator trust lag. This asymmetry has clear practical implications (Mst Shurovi, 2026; Shima Ali et al., 2026): the greatest gains in reliability are likely to come not from further improving already-strong modeling but from strengthening sensor validation, deepening model interpretability, and investing in the operator engagement that converts predictions into trusted maintenance decisions (Carvalho et al., 2019). This interpretation is robust across the descriptive, correlational, regression, maturity, and priority analyses, which consistently identify the same set of improvement areas (Syed Nurul, 2026).

CONCLUSION

This study set out to assess how an XAI digital twin framework, integrating physics-informed digital twin modeling with interpretable machine learning, influences predictive failure detection and lifecycle management performance in U.S. drinking water treatment infrastructure. Using a quantitative, cross-sectional, case-based design and data from 138 valid respondents, the study found that all six hypothesized relationships were supported and that the framework's constructs together explained 70.2% of the variance in lifecycle management performance. Framework design quality emerged as the strongest predictor, underscoring that the value of digital twins and machine learning is realized through coherent integration rather than isolated technique. Digital twin modeling and XAI

interpretability were confirmed as central technical capabilities, while sensor validation, predictive detection accuracy, and operator engagement were identified as significant but comparatively less mature contributors.

The study concludes that an XAI digital twin framework offers a substantial improvement over reactive, schedule-based, and opaque condition monitoring for drinking water utilities, because it extends the reach and transparency of prediction, links model outputs to physical understanding, and supports lifecycle decision-making. However, the study also concludes that the reliability of such a framework is fundamentally socio-technical: its performance depends not only on the quality of digital twin modeling and machine learning but also on the rigor of sensor validation and the strength of the interpretive and institutional arrangements that convert predictions into trusted action. The consistent identification of sensor validation, model interpretability, and operator trust as the highest-priority improvement areas points to where future effort should be concentrated. In sum, the framework provides a promising and empirically supported basis for strengthening predictive asset management in the water sector, provided that its technical strengths are matched by corresponding investment in data quality, explainability, and operator engagement.

RECOMMENDATIONS

Based on the findings, several recommendations are offered for utilities, engineers, and regulators engaged in predictive water-asset management. First, organizations should strengthen sensor validation and calibration by systematically checking sensor drift, cross-validating readings, and documenting data-quality procedures, since sensor validation recorded the largest improvement gap. Second, utilities should invest in model interpretability, prioritizing methods that make predictions explainable to operators and regulators rather than optimizing accuracy alone, so that predictions can be trusted, justified, and acted upon. Third, utilities should develop operator-engagement and training practices that build trust in model outputs and integrate predictions into maintenance workflows, addressing the persistent gap between predictive capability and human acceptance. Fourth, organizations should establish routine model-updating and audit-readiness procedures, including consistent data governance and documentation, to ensure that twins and models remain current and defensible over time. Fifth, utilities should prioritize the integration of membrane, pump, and distribution-asset condition within a common digital twin, given the high importance and moderate effectiveness of these areas. Finally, cross-disciplinary coordination among operators, engineers, data scientists, and asset managers should be formalized, consistent with the socio-technical finding that reliable asset management depends on collaboration across professional groups.

LIMITATIONS OF THE STUDY

This study has several limitations that should be considered when interpreting its findings. First, the cross-sectional design captured perceptions at a single point in time and therefore cannot establish causal direction or track how framework performance changes as digital twin adoption matures; a longitudinal design would allow stronger causal inference. Second, the study relied on a structured, perception-based questionnaire completed by professionals, so the results reflect informed expert assessment rather than direct measurement of prediction accuracy or avoided failures; future work could pair perception data with technical performance metrics such as validated detection rates and remaining-useful-life accuracy. Third, the purposive sampling strategy, while appropriate for reaching knowledgeable respondents, limits the generalizability of the findings to the broader population and to utilities not represented in the sample. Fourth, the framework and its constructs were assessed at a general level across diverse utility contexts, and the relative importance of specific assets and models may vary considerably between small and large systems, and between surface-water and groundwater sources. Finally, the study did not directly evaluate the cost, data-availability, cybersecurity, or workforce-capacity constraints that may affect real-world implementation, which represent important avenues for future research.

REFERENCES

- [1]. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
- [2]. Amir, R. (2026). AI-Enabled Predictive Maintenance for Fire Alarm and Smoke Management Systems: A Systematic Review of Models. *American Journal of Interdisciplinary Studies*, 7(01), 496-535. <https://doi.org/10.63125/q3aw4y77>
- [3]. ASCE. (2021). 2021 report card for America's infrastructure: Drinking water. American Society of Civil Engineers.

- <https://infrastructurereportcard.org/cat-item/drinking-water-infrastructure/>
- [4]. Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bannetot, A., Tabik, S., Barbado, A., ... Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- [5]. Bartos, M., & Kerkez, B. (2021). Pipedream: An interactive digital twin model for natural and urban water systems. *Environmental Modelling & Software*, 144, 105124. <https://doi.org/10.1016/j.envsoft.2021.105124>
- [6]. Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. da P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137, 106024. <https://doi.org/10.1016/j.cie.2019.106024>
- [7]. Conejos Fuertes, P., Martínez Alzamora, F., Hervás Carot, M., & Alonso Campos, J. C. (2020). Building and exploiting a digital twin for the management of drinking water distribution networks. *Urban Water Journal*, 17(8), 704–713. <https://doi.org/10.2166/hydro.2020.021>
- [8]. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- [9]. Eggimann, S., Mutzner, L., Wani, O., Schneider, M. Y., Spuhler, D., Moy de Vitry, M., ... Maurer, M. (2017). The potential of knowing more: A review of data-driven urban water management. *Environmental Science & Technology*, 51(5), 2538–2553. <https://doi.org/10.1021/acs.est.6b06135>
- [10]. El-Diraby, T. E., Kinawy, S., & Pirayonesi, S. M. (2017). A comprehensive review of approaches used by Ontario municipalities to develop road asset management plans. *Journal of Infrastructure Systems*, 23(4). [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000345](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000345)
- [11]. Folkman, S. (2018). Water main break rates in the USA and Canada: A comprehensive study. Utah State University Buried Structures Laboratory. https://digitalcommons.usu.edu/mae_facpub/174/
- [12]. Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. *IEEE Access*, 8, 108952–108971. <https://doi.org/10.1109/ACCESS.2020.2998358>
- [13]. Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary perspectives on complex systems* (pp. 85–113). Springer. https://doi.org/10.1007/978-3-319-38756-7_4
- [14]. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), 1–42. <https://doi.org/10.1145/3236009>
- [15]. Halfawy, M. R. (2008). Integration of municipal infrastructure asset management processes: Challenges and solutions. *Journal of Computing in Civil Engineering*, 22(4), 216–229. [https://doi.org/10.1061/\(ASCE\)0887-3801\(2008\)22:4\(216\)](https://doi.org/10.1061/(ASCE)0887-3801(2008)22:4(216))
- [16]. Hill, D. J., & Minsker, B. S. (2009). Anomaly detection in streaming environmental sensor data: A data-driven modeling approach. *Environmental Modelling & Software*, 25(9), 1014–1022. <https://doi.org/10.1016/j.envsoft.2009.02.011>
- [17]. Hollnagel, E., Woods, D. D., & Leveson, N. (Eds.). (2006). *Resilience engineering: Concepts and precepts*. Ashgate. <https://doi.org/10.1201/9781315605685>
- [18]. Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the digital twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, 29, 36–52. <https://doi.org/10.1016/j.cirpj.2020.02.002>
- [19]. Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3(6), 422–440. <https://doi.org/10.1038/s42254-021-00314-5>
- [20]. Kazi Rakib Hasan, S., & Uddin, H. M. M. (2022). Scalable AI For Project Portfolio Management: A Mixed-Methods Study Combining Distributed Computing Benchmarks. *Review of Applied Science and Technology*, 1(04), 375–410. <https://doi.org/10.63125/0kk4wf20>
- [21]. Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11), 1016–1022. <https://doi.org/10.1016/j.ifacol.2018.08.474>
- [22]. Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & Siegel, D. (2014). Prognostics and health management design for rotary machinery systems: Reviews, methodology and applications. *Mechanical Systems and Signal Processing*, 42(1–2), 314–334. <https://doi.org/10.1016/j.ymssp.2013.06.004>
- [23]. Lockey, S., Gillespie, N., Holm, D., & Someh, I. A. (2021). A review of trust in artificial intelligence: Challenges, vulnerabilities and future directions. *Proceedings of the 54th Hawaii International Conference on System Sciences*, 5463–5472. <https://doi.org/10.24251/HICSS.2021.664>
- [24]. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774. <https://doi.org/10.48550/arXiv.1705.07874>
- [25]. Mahmudul, H., & Sadia, Z. (2026). Safeguarding Financial Records: A Cybersecurity-Driven Management Framework. *International Journal of Scientific Interdisciplinary Research*, 7(1), 493–526. <https://doi.org/10.63125/zch4s169>
- [26]. Marlow, D. R., Beale, D. J., & Burn, S. (2010). A pathway to a more sustainable water sector: Sustainability-based asset management. *Water Science and Technology*, 61(5), 1245–1255. <https://doi.org/10.1016/j.resconrec.2010.03.008>
- [27]. Md Arif Uz, Z., Sharmin, S., & Md Abdur, R. (2025). Factors Influencing Behavioural Intention to Use he Innovative Open And Distance Learning Systems With The Role Of Continuance Intention. *American Journal of Scholarly Research and Innovation*, 4(01), 202–220. <https://doi.org/10.63125/pbjxp014>
- [28]. Md Tohidul, I. (2023). ERP-Integrated Financial Analytics Systems for Corporate Financial Reporting, Cash Flow Analysis, and Strategic Investment Planning. *American Journal of Advanced Technology and Engineering Solutions*, 3(04), 87–128. <https://doi.org/10.63125/qhk96h35>
- [29]. Md. Abdur, R., & Iftekhar, A. (2021). Customer Retention Forecasting in Mobile Wallet Services Using Neural Networks: A Comparative Quantitative Study. *International Journal of Business and Economics Insights*, 1(4), 70–102.

- <https://doi.org/10.63125/dyrpc387>
- [30]. Mohammad Badrul, A., & Md. Mominul, H. (2023). Machine Learning-Based Assessment of Environmental and Public Health Risks in Sewerage Sludge Management Systems. *American Journal of Advanced Technology and Engineering Solutions*, 3(02), 33-77. <https://doi.org/10.63125/y3yfxa92>
- [31]. Mohammad Shoriful Hossan, S. (2024). Development of A Predictive Maintenance Model for Industrial Motor Drives and VFD Systems Using PLC/DCS Operational Telemetry. *American Journal of Scholarly Research and Innovation*, 3(01), 198-242. <https://doi.org/10.63125/pv1js496>
- [32]. Mohammad Shoriful Hossan, S. (2025). Development Of an ARC-Flash Hazard Mitigation Framework for Brownfield Industrial Power Distribution Retrofits in U.S. Manufacturing. *American Journal of Advanced Technology and Engineering Solutions*, 1(02), 262-303. <https://doi.org/10.63125/vhqjgm79>
- [33]. Mohammad Shoriful Hossan, S., & Adar, C. (2023). SIL-Compliant Safety Instrumented System Design for Oil & Gas Hazardous-Area Operations Using IEC 61508 / IEC 61511 Frameworks. *Journal of Sustainable Development and Policy*, 2(02), 43-85. <https://doi.org/10.63125/cb3d3w81>
- [34]. Mst Shurovi, A. (2026). AI-Driven Accounts Payable and Receivable Automation for Operational Risk Mitigation in U.S. SMEs: A Systematic Review (2018–2026). *American Journal of Interdisciplinary Studies*, 7(01), 536-577. <https://doi.org/10.63125/x1yzwc45>
- [35]. Muhammad Mohiul, I., & Mohammad Badrul, A. (2022). Predictive Analytics for Optimizing Sewerage Sludge Treatment Efficiency and Resource Recovery in Dhaka City. *Journal of Sustainable Development and Policy*, 1(04), 158-200. <https://doi.org/10.63125/t9zdp458>
- [36]. Mukut Kanti, B. (2025). Lifecycle Carbon Reduction and Net Zero Pathway Modeling Using LEED, EDGE, and SBTI Frameworks in Developing Economy Industrial Buildings. *American Journal of Interdisciplinary Studies*, 6(3), 306-350. <https://doi.org/10.63125/tnneth75>
- [37]. Mukut Kanti, B., & Rony, S. (2022). Quantitative Risk Assessment and Safety Performance Benchmarking for Structural Integrity Management in Retail Supply Chain Infrastructure. *American Journal of Data Science and Analytics*, 3(12), 42-88. <https://doi.org/10.63125/n88vth90>
- [38]. Nam, S.-H., Kim, S., & Kim, J. (2019). Development of a fault detection framework for drinking water treatment processes. *Water Research*, 165, 114924. <https://doi.org/10.1016/j.watres.2019.114924>
- [39]. Pesantez, J. E., Alghamdi, F., Sabu, S., Mahinthakumar, G., & Berglund, E. Z. (2022). Using a digital twin to explore water infrastructure impacts during the COVID-19 pandemic. *Sustainable Cities and Society*, 77, 103520. <https://doi.org/10.1016/j.envsoft.2022.105442>
- [40]. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- [41]. Rajani, B., & Kleiner, Y. (2001). Comprehensive review of structural deterioration of water mains: Physically based models. *Urban Water*, 3(3), 151–164. <https://doi.org/10.1080/10286600108830280>
- [42]. Ramos, H. M., McNabola, A., López-Jiménez, P. A., & Pérez-Sánchez, M. (2023). Smart water management towards future water sustainable networks. *Water*, 15(6), 1129. <https://doi.org/10.3390/w15061129>
- [43]. Ran, Y., Zhou, X., Lin, P., Wen, Y., & Deng, R. (2019). A survey of predictive maintenance: Systems, purposes and approaches. *arXiv preprint arXiv:1912.07383*. <https://doi.org/10.48550/arXiv.1912.07383>
- [44]. Rasheed, A., San, O., & Kvamsdal, T. (2020). Digital twin: Values, challenges and enablers from a modeling perspective. *IEEE Access*, 8, 21980–22012. <https://doi.org/10.1109/ACCESS.2020.2970143>
- [45]. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why should I trust you?”: Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- [46]. Risha, A., & Kazi Mohammad Khalid, A. (2023). A Meta-Analysis of AI-Driven Geospatial Analytics for Predictive Maintenance of Critical Infrastructure in Developing Economies. *International Journal of Scientific Interdisciplinary Research*, 4(4), 375–412. <https://doi.org/10.63125/rayrex49>
- [47]. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215. <https://doi.org/10.1038/s42256-019-0048-x>
- [48]. Sadia, Z., Mohammad Ariful, I., Sumyta, H., & Hosne Ara, M. (2022). A Review of AI-Powered Data Visualization in Enterprise Reporting: Dashboard Design And Interactive Analytics. *American Journal of Advanced Technology and Engineering Solutions*, 2(01), 32-54. <https://doi.org/10.63125/gabst658>
- [49]. Shima Ali, S., Musomi, K., & Sonya, G. (2024). Integrated Infectious Disease Preparedness: Multilingual Health Communication Strategies for Public Health Risk Mitigation, Disease Intervention and Prevention for Culturally Diverse Population in United States. *American Journal of Data Science and Analytics*, 5(10), 01-47. <https://doi.org/10.63125/ehk4mt02>
- [50]. Shima Ali, S., Sonya, G., & Mahbubul, I. (2025). Telemedicine And Remote Patient Monitoring for Preventing Risk and Chronic Disease Management in Underserved Communities. *American Journal of Health and Medical Sciences*, 6(04), 01–42. <https://doi.org/10.63125/y0e9dc38>
- [51]. Shima Ali, S., Sonya, G., & Mahbubul, I. (2026). STI and HIV Prevention Through Public Health Digital Twins: A Framework for Personalized Prevention and Adaptive Disease Intervention. *American Journal of Scholarly Research and Innovation*, 5(01), 284–326. <https://doi.org/10.63125/rjgna882>
- [52]. Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1145/3419764>
- [53]. Syed Nurul, I. (2026). A Machine Learning Approach to Intelligent Streetlight Control for Sustainable Energy

- Management in Smart Cities. *American Journal of Advanced Technology and Engineering Solutions*, 6(01), 650-689. <https://doi.org/10.63125/gxwn1g63>
- [54]. Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2019). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, 15(4), 2405–2415. <https://doi.org/10.1109/TII.2018.2873186>
- [55]. Trist, E. (1981). The evolution of socio-technical systems: A conceptual framework and an action research program. Ontario Quality of Working Life Centre. <https://doi.org/10.4324/9781315227979>
- [56]. U.S. EPA. (2020). Drinking water infrastructure needs survey and assessment: Seventh report to Congress. United States Environmental Protection Agency. <https://www.epa.gov/dwsrf/drinking-water-infrastructure-needs-survey-and-assessment>
- [57]. Weick, K. E., & Sutcliffe, K. M. (2007). *Managing the unexpected: Resilient performance in an age of uncertainty* (2nd ed.). Jossey-Bass. <https://www.wiley.com/en-us/Managing+the+Unexpected%3A+Resilient+Performance+in+an+Age+of+Uncertainty%2C+2nd+Edition-p-9780787996499>
- [58]. 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. <https://doi.org/10.1145/3514228>
- [59]. Zaman, S. (2024). A Systematic Review of ERP And CRM Integration For Sustainable Business And Data Management in Logistics And Supply Chain Industry. *Frontiers in Applied Engineering and Technology*, 1(01), 204–221. <https://doi.org/10.70937/faet.v1i01.36>
- [60]. Zhang, W., Yang, D., & Wang, H. (2019). Data-driven methods for predictive maintenance of industrial equipment: A survey. *IEEE Systems Journal*, 13(3), 2213–2227. <https://doi.org/10.1109/ACCESS.2019.2925008>
- [61]. Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the Industry 4.0: A systematic literature review. *Computers & Industrial Engineering*, 150, 106889. <https://doi.org/10.1016/j.cie.2020.106889>