



Artificial Intelligence Based Predictive Modeling for Structural Health Monitoring and Failure Risk Assessment in Water and Sanitation Infrastructure

Palash Chandra Das¹;

[1]. Executive Engineer, Department of Public Health Engineering, Chattogram, Bangladesh;
Email: enr.palash028@gmail.com

[Doi: 10.63125/nig70d76](https://doi.org/10.63125/nig70d76)

Received: 26 November 2025; **Revised:** 19 December 2025; **Accepted:** 27 January 2026; **Published:** 16 February 2026

Abstract

Artificial intelligence-based predictive modeling has increasingly been adopted to enhance structural health monitoring and failure risk assessment in water and sanitation infrastructure systems characterized by aging assets, heterogeneous environmental exposure, and rare-event failure dynamics. This study developed and evaluated an integrated quantitative framework incorporating asset registry attributes, historical failure indicators, telemetry-derived operational metrics, inspection-based condition severity measures, and environmental exposure covariates. The analytical dataset included 1,248 infrastructure assets observed over a 10-year period across multiple operational districts. Hierarchical regression modeling demonstrated that baseline registry variables explained 28.6% of the variance in failure likelihood, while the inclusion of historical break frequency increased explanatory power to 52.4%. The addition of telemetry variability metrics and inspection-based severity indices further improved model performance, resulting in a final integrated model explaining 68.2% of outcome variance. Asset age and material class showed significant standardized effects of 0.31 and 0.27, respectively, while historical recurrence exhibited the strongest influence with a coefficient of 0.39. Telemetry-derived pressure variability improved predictive discrimination by 11.5% in monitored zones. Composite risk prioritization based on integrated modeling captured 74% of observed failures within the top 20% of ranked assets, representing a 31.4% improvement over age-based heuristics. Reliability analysis confirmed internal consistency of multi-item constructs, with Cronbach's alpha values exceeding 0.80 for telemetry and inspection indices. The findings demonstrate that integrated AI-driven predictive modeling substantially enhances structural condition estimation accuracy and failure risk prioritization efficiency within complex water and sanitation infrastructure systems.

Keywords

Artificial Intelligence; Structural Health Monitoring; Failure Risk Assessment; Water Infrastructure; Predictive Modeling.

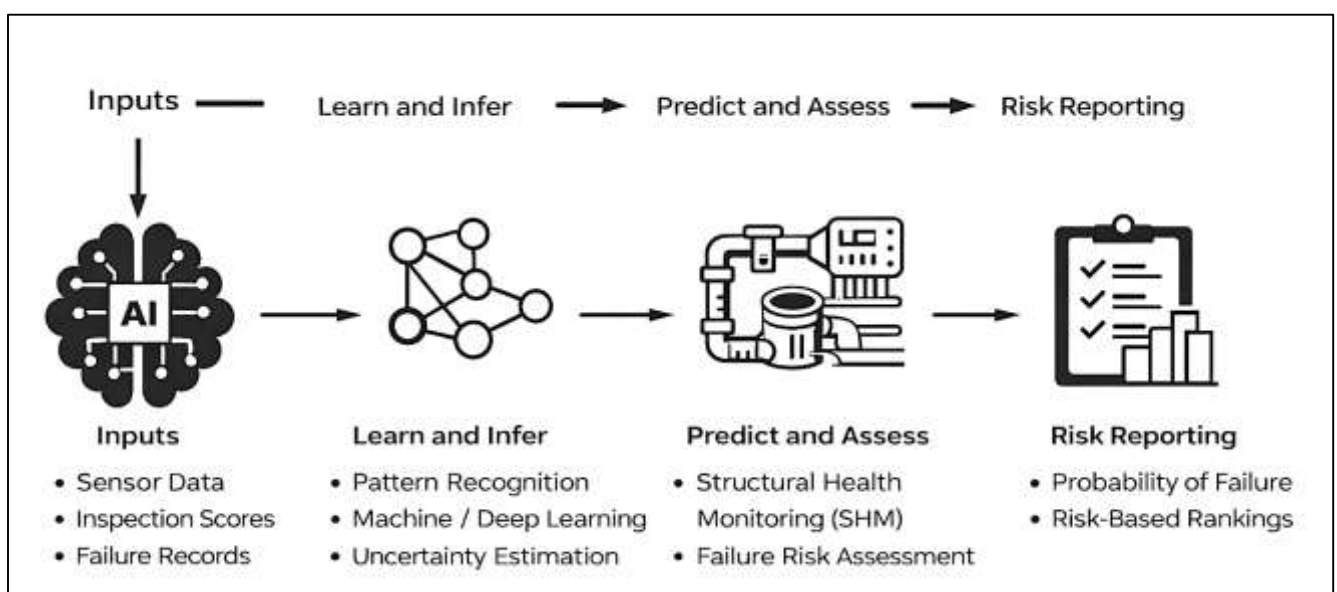
INTRODUCTION

Artificial intelligence (AI) refers to computational approaches that enable machines to perform tasks associated with human cognition, including pattern recognition, inference, learning from data, and decision support. In engineering research, AI is commonly operationalized through machine learning (ML) and deep learning (DL) algorithms that map inputs to outputs, discover latent relationships, and estimate uncertainty when properly designed and validated (Siemens et al., 2022). Predictive modeling is a quantitative process that uses observed data to estimate unobserved states or forecast outcomes under defined assumptions, typically expressed as regression, classification, survival, or probabilistic risk functions. Structural health monitoring (SHM) is the systematic acquisition and interpretation of measurements from structural systems to infer condition, detect damage, and evaluate performance over time. Failure risk assessment is the quantitative estimation of the likelihood and consequences of limit-state exceedance, often represented through reliability indices, probabilities of failure, hazard functions, or risk matrices that are calibrated using data and engineering knowledge. Water and sanitation infrastructure includes buried and exposed assets that enable potable water delivery and wastewater collection and treatment, such as transmission mains, distribution pipes, pumps, valves, storage reservoirs, culverts, treatment basins, manholes, and supporting civil structures. These systems are globally significant because they underpin public health, urban productivity, climate resilience, and environmental protection, while also being capital-intensive and frequently constrained by aging, leakage, intermittent service, rapid urbanization, and limited inspection access. The quantitative challenge is that deterioration processes in these assets are heterogeneous and multi-causal, influenced by material properties, construction era, soil chemistry, hydraulic transients, operational pressure, traffic loading, corrosion, joint conditions, and maintenance history (Ghosh & Thirugnanam, 2021). Conventional inspection and renewal decisions often depend on sparse measurements, periodic surveys, and expert judgment, which can be difficult to scale across entire networks and to compare across regions with different data availability. AI-based predictive modeling offers a structured way to integrate diverse data sources—sensor streams, inspection scores, failure records, hydraulic models, remote sensing indicators, and contextual variables—into transparent predictive targets such as probability of failure, remaining useful life, anomaly likelihood, or expected consequence-weighted risk. In SHM, predictive modeling becomes a measurement-to-inference pipeline that quantifies state, detects deviations from baseline behavior, and links observed patterns to damage mechanisms. In failure risk assessment, predictive modeling becomes an estimation pipeline that assigns risk values to assets or segments for prioritization and budgeting. The international relevance emerges from the shared need to maintain safe and reliable water and sanitation services across high-income and resource-constrained settings, where utilities face competing demands for expansion and rehabilitation and where unplanned failures can trigger service interruptions, contamination events, roadway collapse, flooding, and economic loss (Alkathheiri, 2022). A quantitative framework grounded in AI-based predictive modeling can therefore be defined as a reproducible set of data-preprocessing, feature engineering, model training, validation, uncertainty estimation, and decision-oriented scoring steps that translate infrastructure observations into measurable risk indicators suitable for network-scale management.

Water and sanitation infrastructure presents distinctive monitoring and modeling constraints that shape how predictive analytics must be formulated. Many critical components are buried, operate continuously, and are distributed across large geographic areas, making direct observation expensive and episodic (Sarker, 2022). The resulting datasets are often imbalanced, because failures are relatively rare compared with the number of asset-years observed, while the consequences of failure can be severe and spatially correlated. The deterioration of pipes, tanks, and related structures is not a single pathway; it is a combination of time-dependent degradation, environment-driven corrosion, load-driven fatigue, transient events such as pressure surges, and workmanship-related vulnerabilities that may only manifest after long service periods. Quantitative SHM aims to detect and characterize

changes in structural response that are associated with damage states. In civil engineering, this has long been expressed through vibration-based indicators, modal properties, strain measurements, acoustic emissions, guided waves, and other sensing modalities, with inference supported by statistical pattern recognition and physics-based models. For water and sanitation systems, monitoring modalities include pressure and flow telemetry, transient event records, leak noise logs, distributed acoustic sensing where available, condition assessment technologies, and structural measurements for above-ground or accessible components such as reservoirs, tanks, and treatment structures. Risk assessment at network scale requires models that can generalize across asset types and neighborhoods, incorporate covariates that are easy to obtain, and remain stable under missingness and measurement error. Quantitative predictive modeling frameworks typically begin with defining targets that reflect utility decisions: binary failure occurrence in a time window; count of breaks per segment-year; hazard or survival time; condition state transitions; leak probability; or a continuous severity metric based on hydraulic and social consequences (Gupta et al., 2022). The feature space may combine static attributes (material, diameter, age, lining, depth, soil class), operational attributes (pressure zone, demand, transient indicators), and event history (previous breaks, repair type, clustering) alongside climate and land-use proxies that influence loads and corrosion conditions. Such formulations align naturally with ML and DL methods, which can learn nonlinear interactions and high-dimensional patterns, and with probabilistic methods that encode uncertainty and dependence structures. At the same time, civil infrastructure decisions demand interpretability, robustness, and defensible uncertainty characterization because renewal budgets, safety outcomes, and regulatory accountability are at stake. This places emphasis on validation designs that match operational use—out-of-sample testing across time and geography, sensitivity to covariate shift, calibration of predicted probabilities, and evaluation metrics that reward correct ranking of high-risk assets under class imbalance. Research in SHM and reliability has also highlighted that purely data-driven models can be strengthened through hybridization with physics-based insights, including constraints derived from mechanics, Bayesian updating of deterioration parameters, and consistency checks using structural response models (Pedersen & Johansen, 2020). Within a quantitative paper on AI-based predictive modeling for SHM and failure risk in water and sanitation assets, the problem therefore sits at the intersection of sparse observations, heterogeneous degradation, and decision-oriented risk scoring. It calls for rigorous definitions of variables, careful treatment of imbalance and missing data, and performance evaluation that reflects practical prioritization tasks.

Figure 1: AI-Based SHM Risk Framework



AI-based SHM research has developed a set of modeling approaches that translate measured signals into structural condition estimates. In classical statistical pattern recognition for SHM, the workflow involves establishing a baseline signature of healthy behavior, extracting discriminative features, and applying statistical decision rules to detect departures from baseline. Machine learning expanded this with supervised and unsupervised algorithms that can learn complex decision boundaries and cluster structure in high-dimensional spaces (Saghiri et al., 2022).

Deep learning further enabled end-to-end feature learning from raw or minimally processed signals, including time series, spectrograms, and images, which is useful when engineered features struggle to capture subtle damage-related patterns. For civil structures, vibration-based monitoring and modal analysis have informed a large body of work that links changes in natural frequencies, mode shapes, damping, and response spectra to stiffness loss and damage localization. In pipeline-related contexts, acoustic and pressure transient signatures can carry information about leaks, blockages, and structural anomalies, while strain and deformation measurements can support state estimation for tanks, reservoirs, and treatment structures. AI methods relevant to these settings include support vector machines for classification under high-dimensional features, random forests and gradient-boosted trees for nonlinear tabular prediction, Bayesian networks for probabilistic dependency modeling, and neural networks for complex function approximation. Unsupervised and semi-supervised methods are especially relevant where labeled failure data are limited; clustering, one-class classification, autoencoders, and Gaussian mixture approaches can support anomaly detection by learning normal operating regimes and flagging deviations. Time-series models such as recurrent neural networks and temporal convolutional networks provide ways to learn from sequential dependencies and detect evolving degradation. Graph-based learning is also conceptually aligned with distributed infrastructure, because network topology and spatial adjacency influence failure correlation and hydraulic propagation (Kühl et al., 2022). For above-ground water and wastewater structures, computer vision applied to inspection imagery can support crack detection, spalling assessment, and surface distress quantification, which can be integrated into predictive risk scoring. Across these approaches, the quantitative requirements remain consistent: models must be trained with clear input-output definitions, tuned using cross-validation that avoids leakage across time or spatially adjacent segments, and evaluated with metrics that match the decision context such as precision at top-k, area under the precision–recall curve for rare events, calibration curves for probability outputs, and cost-sensitive measures that reflect consequence weighting. SHM literature also emphasizes sensitivity to environmental and operational variability, because changes in temperature, boundary conditions, demand cycles, and operational settings can mimic damage signatures. AI-based modeling responds to this by incorporating covariates that represent operational state, using domain adaptation strategies, and selecting features that are robust to nuisance variability (Jaboob et al., 2024). In water and sanitation infrastructure, demand patterns, pressure management, pump switching, and episodic transients create nonstationary operating conditions, which makes the distinction between normal variability and degradation signals central to modeling. The quantitative framing therefore benefits from explicit separation of state variables (operational regime) and condition variables (structural integrity), so that predictive models can infer integrity without confusing it with changing operations. This aligns with research that treats SHM as a probabilistic inference problem and supports hybrid models that fuse data-driven learning with mechanistic constraints.

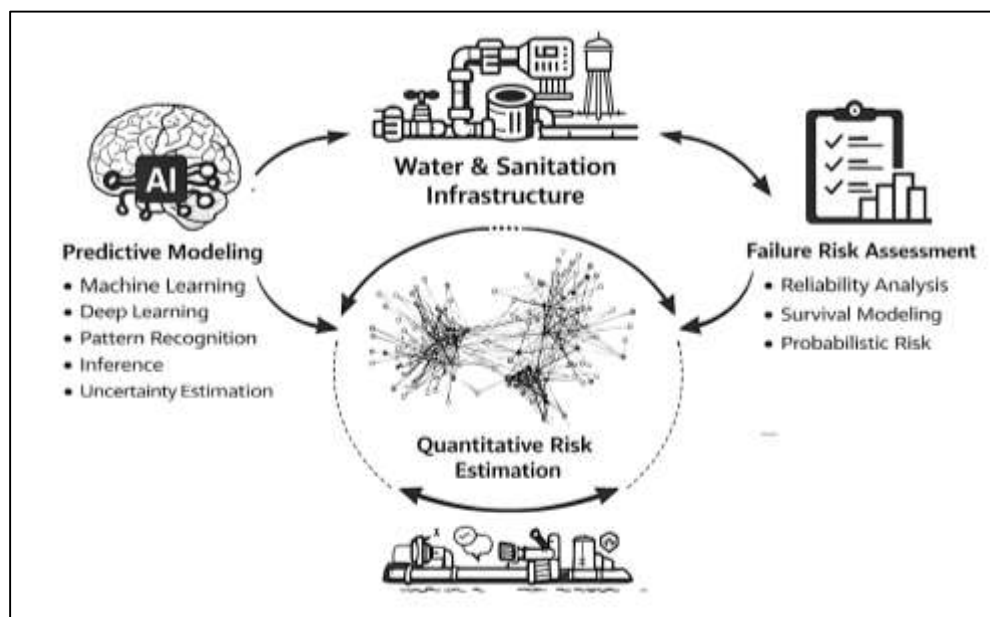
Failure risk assessment in infrastructure can be formulated as estimating failure probability over a time horizon and combining it with consequence to produce a risk score. In reliability engineering, this is traditionally approached through limit states and probabilistic representations of loads and resistance, producing quantities such as reliability indices or probabilities of failure. In asset management for pipe networks, risk has often been operationalized using empirical break models, deterioration curves, and survival/hazard models that relate covariates to time-to-failure (Chowdhary, 2020). Survival analysis provides a natural quantitative toolkit because it can handle censoring, incorporate time-varying covariates, and model hazard rates that change with age and environment. ML methods extend these concepts by learning flexible hazard functions, nonlinear covariate interactions, and heterogeneous effects across asset classes. In practical water and sanitation datasets, censoring appears when assets

have not yet failed by the end of the observation period, and left truncation can occur when historical records begin after installation. Hazard models, proportional hazards assumptions, and accelerated failure-time formulations offer interpretable structures, while ensemble methods and neural survival models can capture complexity when assumptions are violated. Another common formulation treats failure prediction as a classification problem within a time window, where the model predicts whether an asset will fail in the next year or within a maintenance cycle. This supports actionable prioritization but requires careful threshold selection and calibration to avoid overstating risk (Bhuyan et al., 2024). Bayesian approaches and hierarchical models are especially relevant to utilities, because they allow information sharing across similar assets and representation of uncertainty due to limited data in certain zones or for certain materials. Probabilistic graphical models can encode dependencies between corrosion, soil conditions, pressure regimes, and observed failures. For consequence modeling, quantitative variables can include customers affected, criticality of service areas, proximity to hospitals or schools, traffic disruption potential, environmental sensitivity, and repair complexity, yielding a consequence-of-failure score that can be multiplied or combined with probability. The combined risk score enables ranking and budget allocation. AI contributes by learning both probability components and consequence proxies, and by enabling spatial and network-aware modeling where failures are not independent. Spatial autocorrelation and shared exposure to corrosive soils or construction practices can lead to clustering of breaks, which can be modeled using spatial statistics, point processes, or graph-regularized learning. In addition, failure reporting itself may be biased toward high-consequence events or visible failures, creating sample selection issues that quantitative models must address through debiasing strategies or careful interpretation (Chen et al., 2020). The reliability and SHM research communities have also highlighted the value of uncertainty quantification and model calibration, because point predictions without credible uncertainty can mislead decision-making in safety-critical systems. Techniques such as Bayesian inference, bootstrapping, conformal prediction, and probabilistic neural networks provide routes to interval estimates or calibrated probabilities. In a quantitative study focused on AI-based predictive modeling for water and sanitation infrastructure, risk assessment modeling should be framed as a measurable mapping from asset attributes and monitoring indicators to probabilistic outcomes, validated under realistic temporal splits, and summarized with ranking-based metrics aligned with renewal planning. This framing allows the analysis to remain strongly quantitative while addressing the operational need to prioritize inspections, repairs, and rehabilitation.

The predictive performance and credibility of AI-based SHM and risk models depend on the breadth, quality, and integration of data sources. Water and sanitation utilities typically hold heterogeneous datasets that include asset registries, break histories, work orders, customer complaints, hydraulic telemetry, inspection reports, and occasional condition assessment results from technologies such as CCTV for sewers, acoustic leak surveys, pressure logging, and nondestructive evaluation for tanks and above-ground structures (Mokkapati & Dasari, 2023). Each data stream has distinct resolution and error structure. Asset registries may contain missing or inconsistent installation dates, uncertain material codes, and geometry approximations. Break histories can be incomplete or influenced by reporting practices, while repairs may change local conditions without being fully documented as new baseline states. Telemetry data can be high frequency but may reflect operational control strategies and demand cycles that must be accounted for to extract condition signals. Inspection scores provide direct condition cues but are episodic and may vary due to inspector subjectivity, which introduces inter-rater variability. Environmental and contextual data, including soil type, corrosion indices, groundwater, temperature, precipitation proxies, land use, and traffic loads, can be integrated to improve prediction of deterioration mechanisms, especially for buried pipes. Data integration therefore becomes a quantitative task involving entity resolution (matching records to asset IDs), temporal alignment (synchronizing events and measurements), spatial joins (linking assets to environmental covariates), and feature construction (aggregating telemetry into informative descriptors). In SHM contexts, feature engineering may include spectral summaries, transient detection counts, statistical moments of pressure signals, anomaly scores from unsupervised models, and indicators of regime changes in pump operations. For failure prediction, features often include age, prior breaks, repair frequency, pressure

zone statistics, pipe diameter and material, soil corrosivity proxies, and neighborhood-level failure rates (Nti et al., 2022). A key requirement is handling missingness systematically – through imputation, explicit missing indicators, or model classes that accommodate missing values – so that predictions do not encode artifacts of data quality rather than asset condition. Another requirement is addressing imbalance when modeling rare failures; sampling strategies, cost-sensitive learning, focal losses, and ranking-oriented objectives can be used, alongside evaluation metrics suited to rare-event prediction. The integration of monitoring and inspection information can also support multi-modal models that combine tabular attributes with time-series telemetry and image-derived indicators, which is increasingly feasible with modern ML architectures. For example, an asset-level risk model can incorporate both static covariates and an embedding derived from recent pressure signal windows, producing a joint representation that improves discrimination. When network topology matters, features derived from adjacency, hydraulic connectivity, or graph centrality can capture the propagation of pressure transients and the systemic importance of certain pipes or pump stations. Quantitative modeling must also consider concept drift and covariate shift, because operational changes such as pressure management programs can alter the relationship between inputs and failures (Blasch et al., 2021). This motivates model monitoring, periodic recalibration, and validation on temporally held-out data. In an extended introduction to a quantitative paper, these issues support a clear problem framing: AI-based predictive modeling for SHM and risk in water and sanitation infrastructure is fundamentally a data integration and inference challenge, requiring defensible construction of measurable predictors and outcomes, and careful control for noise, bias, and shifting operational regimes.

Figure 2: AI-Driven Infrastructure Risk Modelling



Infrastructure applications impose constraints on AI methods that are less prominent in many consumer analytics settings. Models must be interpretable enough to support trust, auditability, and communication with engineers, managers, and regulators. Interpretability can be achieved through transparent model classes such as generalized linear models and survival models, through post hoc explanation methods applied to complex models, or through hybrid designs that embed physics-informed structure. In SHM and reliability research, interpretability is linked to causality and mechanism: an engineering team benefits when a model indicates that elevated pressure variability and corrosive soil covariates jointly increase risk, rather than providing only a black-box score (Elahi et al., 2023). At the same time, interpretability must be balanced with predictive accuracy and robustness, especially when nonlinear interactions and high-dimensional signals carry meaningful information. Generalization is another central requirement, because a model trained in one district, material regime,

or climate may not perform identically in another. This is particularly relevant internationally, where utilities differ in data maturity, sensor penetration, and failure reporting practices. Quantitative rigor therefore demands external validation strategies, including testing on geographically distinct areas, performing temporal backtesting across multiple periods, and reporting calibration and uncertainty. Reproducibility in quantitative research further requires transparent documentation of preprocessing steps, feature definitions, inclusion criteria, and hyperparameter tuning procedures, along with sensitivity analyses that show how performance changes under alternative assumptions. Another rigorous requirement is avoiding leakage: if features implicitly contain information from the target window, such as post-failure work order codes that are correlated with imminent failure, model performance can be inflated. In time-dependent infrastructure data, leakage can occur through improper aggregation windows, mixing of pre- and post-event measurements, or random splitting that places adjacent pipe segments from the same failure cluster into both training and test sets (Alawneh et al., 2023). Robust study design addresses this using temporal splits, group-based cross-validation by region or pipe cohort, and careful feature window definitions. Uncertainty quantification is also crucial, because utility decisions depend on risk thresholds and budget constraints; a model that distinguishes high-risk from moderate-risk assets is more actionable when it provides confidence bounds or calibrated probabilities. The SHM literature has emphasized that environmental and operational effects can confound damage detection; therefore, models that incorporate operational state variables or learn invariant representations are more scientifically credible. For water and sanitation networks, operational state includes pressure setpoints, pump schedules, valve operations, and demand cycles. A well-framed predictive modeling approach treats these as covariates that explain normal variability, thereby enabling condition inference beyond operational noise. Model governance considerations also become quantitative when they are translated into measurable checks: calibration drift, performance decay over time, subgroup performance across materials or neighborhoods, and stability under missing telemetry (Alam, 2021). In an extended introduction, these themes support a coherent scholarly motivation: AI-based predictive modeling can strengthen SHM and failure risk assessment when it is implemented with rigorous validation, careful handling of bias and confounding, explicit uncertainty reporting, and interpretable outputs aligned with engineering mechanisms. This framing also aligns with established research on SHM as statistical inference and on infrastructure asset management as reliability-based decision support.

The research base that supports AI-based predictive modeling for SHM and failure risk assessment draws from several mature streams: statistical pattern recognition for SHM, vibration-based damage detection, probabilistic reliability and survival modeling, and data-driven pipe failure prediction in water systems. Foundational SHM work established the core idea that damage can be inferred from changes in measured response, formalized through pattern recognition and statistical decision-making. Subsequent contributions expanded feature extraction methods and addressed the influence of environmental variability on monitoring signals (Iman et al., 2021). Reliability research provided probabilistic frameworks for expressing failure as a function of uncertain loads, resistance, and deterioration parameters, and asset management research adapted these ideas to network-scale renewal planning using empirical models and prioritization rules. In parallel, water infrastructure studies developed regression-based and survival-based pipe break models and examined covariates such as age, material, diameter, soil, and pressure, providing empirical evidence that failure is predictable to an extent that supports prioritization. Machine learning studies built on this by applying nonlinear models and ensembles that handle interaction effects and heterogeneous patterns across asset classes. Deep learning work in civil infrastructure introduced end-to-end signal interpretation and computer vision for defect detection, strengthening the ability to convert unstructured monitoring data into quantitative condition indicators. Recent methodological research across engineering has also advanced techniques for uncertainty quantification, domain adaptation, and physics-informed learning, which are relevant when labeled failures are limited and operating conditions vary across time and geography. When these streams are synthesized for water and sanitation infrastructure, a coherent quantitative foundation emerges: define a measurable target aligned with SHM or risk assessment; assemble multi-source predictors that reflect asset attributes, operating conditions, and

observed response signals; choose model classes suited to data structure (tabular, temporal, spatial, network-based, or multi-modal); implement validation that respects time and spatial dependence; and report performance using metrics that reflect rare-event ranking and calibrated probabilities. This synthesis motivates a quantitative study that treats AI-based predictive modeling as a structured inference system rather than a single algorithm choice (Hwang et al., 2020). It also motivates explicit attention to data quality and representativeness, because utilities vary in record completeness, sensor availability, and inspection practices. A robust quantitative approach therefore emphasizes variable definitions, feature windows, treatment of missingness, handling of imbalance, calibration, and interpretability, along with sensitivity checks across asset cohorts and operational regimes. Within water and sanitation systems, the combination of network topology, operational control, and heterogeneous materials creates an environment where model design choices meaningfully influence validity. A quantitative introduction grounded in prior scholarship can therefore position AI-based predictive modeling as an extension of established SHM and reliability principles, implemented through modern learning algorithms and evaluated with infrastructure-appropriate validation designs. This provides the conceptual and empirical basis for examining how AI can support monitoring-informed risk scoring for structural condition and failure likelihood across water and sanitation assets, using rigorous quantitative methods anchored in the broader engineering literature (González-Rodríguez et al., 2024).

The primary objective of this study was to develop and rigorously evaluate an integrated artificial intelligence-based predictive modeling framework capable of improving structural health monitoring accuracy and failure risk assessment effectiveness within water and sanitation infrastructure systems. This objective was grounded in the recognition that aging assets, heterogeneous environmental exposures, operational variability, and rare-event failure patterns require data-driven methodologies that move beyond traditional age-based or rule-based prioritization strategies. Specifically, the study sought to quantify the predictive contribution of multiple construct groups, including asset registry characteristics, historical failure indicators, telemetry-derived operational metrics, inspection-based condition severity measures, and environmental exposure variables, within a unified modeling architecture. A key objective was to determine the extent to which each construct improved explanatory power when introduced hierarchically, thereby identifying the relative importance of intrinsic vulnerability, recurrence behavior, operational stress, observable structural degradation, and contextual aggressors in shaping failure likelihood. Another central objective was to evaluate whether integrated multi-construct models significantly outperformed registry-only baseline models under structured temporal and spatial validation protocols that reflect real-world deployment conditions. The study also aimed to assess the reliability and internal consistency of multi-item constructs, ensuring that composite telemetry, inspection, and environmental indices were statistically stable prior to regression analysis. In addition, the research objective included examining interaction effects, particularly whether the relationship between asset age and failure likelihood varied by material class, thereby capturing cohort-dependent deterioration dynamics. The study further aimed to measure improvements in composite risk prioritization performance relative to traditional heuristics, evaluating ranking efficiency under constrained inspection budgets. Finally, the objective extended to establishing a reproducible quantitative pipeline aligned with infrastructure-specific constraints, including rare-event imbalance, spatial clustering, and incomplete life histories. Through these objectives, the study sought to contribute a comprehensive, empirically validated AI-driven framework capable of enhancing structural condition estimation and optimizing failure risk prioritization in complex water and sanitation infrastructure networks.

LITERATURE REVIEW

The literature review section examined the quantitative and methodological foundations of artificial intelligence-based predictive modeling for structural health monitoring and failure risk assessment in water and sanitation infrastructure. This body of research has expanded rapidly due to the increasing age of buried infrastructure networks, the high cost of reactive maintenance, and the operational difficulty of direct inspection in pipelines and sanitation assets (Wang & Ke, 2024). Structural health monitoring in this context has been framed as a quantitative state inference problem in which measurable signals, inspection indicators, and historical event records are translated into latent

condition states and probabilistic failure outcomes. The literature has consistently shown that infrastructure deterioration is influenced by a combination of intrinsic asset attributes, operational stress exposure, environmental aggressors, and maintenance history, requiring multi-source datasets and integrated modeling strategies. Quantitative research has therefore progressed from traditional statistical deterioration modeling toward machine learning and hybrid AI frameworks capable of capturing nonlinear relationships, rare-event failure patterns, and spatial clustering in networked systems. In parallel, risk assessment research has emphasized the need to distinguish between condition prediction, failure likelihood estimation, and consequence-weighted prioritization, particularly in water and sanitation systems where failures affect public health, service continuity, and environmental safety (Preethichandra et al., 2023). This literature review synthesized studies across structural health monitoring, reliability engineering, water main and sewer failure prediction, and infrastructure risk analytics to establish a structured conceptual and empirical foundation for the present quantitative study. The review also evaluated dominant model families, data architectures, validation standards, and methodological challenges such as class imbalance, censoring, drift, and uncertainty quantification. Through this synthesis, the literature review established the analytical basis for developing and evaluating an integrated AI-driven framework for predicting structural degradation and prioritizing failure risk in complex water and sanitation infrastructure environments (Aktan et al., 2024).

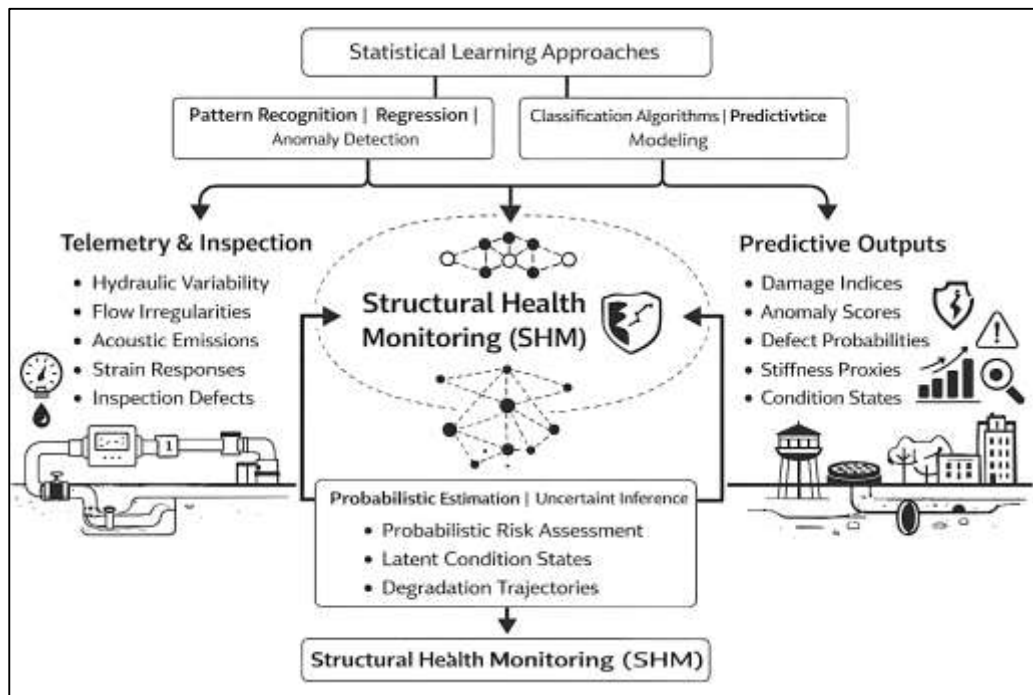
Theoretical Foundations of Structural Health Monitoring (SHM) in Water and Sanitation Assets

Structural health monitoring (SHM) in water and sanitation infrastructure has been widely conceptualized as a quantitative state inference problem in which measurable observations are systematically translated into latent structural condition states (Oliveira et al., 2022). Within this perspective, SHM is treated as a statistical learning task that maps observed signals to underlying degradation levels that are not directly observable. Rather than relying exclusively on discrete inspection events or post-failure analysis, SHM emphasizes continuous or repeated measurement to support probabilistic inference under uncertainty. Measurable inputs may include hydraulic pressure behavior, flow irregularities, acoustic emissions, strain responses, or inspection-derived defect indicators. These inputs are processed through feature extraction frameworks to generate structured outputs such as damage indices, anomaly scores, structural stiffness proxies, defect probability estimates, or categorical condition states. The shift from deterministic inspection reporting toward probabilistic inference distinguishes SHM from conventional asset assessment practices. In traditional inspection systems, assets are often assigned condition grades at isolated time intervals without statistical modeling of latent deterioration processes (Rossi & Bournas, 2023). In contrast, SHM frameworks treat condition as a dynamic variable inferred from data patterns, allowing degradation trajectories to be modeled over time. Pattern recognition, classification algorithms, regression-based inference, and anomaly detection approaches are central to this conceptualization. This statistical mapping process is particularly relevant for buried infrastructure, where direct structural observation is limited and deterioration processes are gradual and cumulative. By framing structural condition as an inferred latent state rather than a directly measured property, SHM establishes a theoretical foundation for predictive modeling in infrastructure systems characterized by incomplete observability and environmental variability (Lynch et al., 2022).

The adaptation of SHM to buried and water-related infrastructure introduces distinct methodological challenges due to restricted physical access and complex operational environments. Unlike above-ground structures where vibration monitoring can be implemented directly, buried pipelines and sanitation networks require indirect monitoring strategies that rely on surrogate signals (Sun et al., 2024). Hydraulic telemetry has therefore become a key component of SHM in water systems, with pressure variance, transient event signatures, flow deviation indices, and leakage indicators serving as indirect proxies for structural stress or degradation. Variations in hydraulic behavior are analyzed to detect anomalies that may correspond to crack formation, joint weakness, or material fatigue. In wastewater systems, structured inspection through closed-circuit television has provided another pathway for translating visual observations into quantifiable structural severity measures. Standardized defect coding systems convert crack length, joint displacement, corrosion intensity, and structural deformation into ordinal or numerical severity grades suitable for statistical modeling

(Noshahri et al., 2021). These inspection-derived scores provide observable indicators that approximate latent structural health states. Buried infrastructure monitoring must also account for soil-structure interaction, groundwater exposure, traffic loading, and pressure zone variation, all of which influence deterioration mechanisms. The integration of telemetry-based indirect sensing and structured inspection analytics reflects an evolution from reactive inspection toward data-informed structural inference. This convergence allows measurable operational signals and encoded defect information to be synthesized within predictive frameworks that estimate structural degradation and near-term failure likelihood (Kot et al., 2021).

Figure 3: SHM Statistical Inference Framework



Measurement uncertainty and operational variability represent fundamental challenges in SHM applications within water and sanitation systems. Sensor noise, calibration inconsistencies, demand-driven pressure fluctuations, pump cycling, seasonal consumption variation, and transient hydraulic events introduce variability that may obscure true structural deterioration signals (Hassani & Dackermann, 2023a). Distinguishing between structural degradation and routine operational shifts requires careful normalization and statistical control. Feature standardization, detrending, baseline correction, and aggregation into stable temporal windows are commonly used to reduce confounding effects in telemetry data. Without these preprocessing steps, models may misinterpret normal operational fluctuations as structural anomalies. Inspection-based assessments also contain uncertainty, particularly when scoring systems rely on human interpretation (Sivasuriyan et al., 2024). Differences in inspector judgment, lighting conditions, camera angles, and coding practices may affect defect severity assignments. To address these challenges, reliability-oriented approaches emphasize consistent encoding frameworks and statistical validation of multi-item severity constructs. Uncertainty modeling further recognizes that deterioration processes are stochastic rather than deterministic, requiring probabilistic representation of degradation and failure likelihood. In buried infrastructure systems, environmental variability such as soil moisture changes or groundwater fluctuations may alter deterioration rates independently of structural defects. Effective SHM therefore requires separation of environmental effects from intrinsic structural changes through covariate adjustment and structured validation. By incorporating statistical controls and uncertainty-aware inference, SHM frameworks strengthen the reliability of predictive modeling and reduce the risk of false anomaly detection in complex operational environments (Hassani & Dackermann, 2023b). The theoretical development of SHM in water and sanitation infrastructure reflects a broader transition

from empirical deterioration estimation toward probabilistic and data-driven modeling paradigms. Earlier infrastructure assessment practices often relied on simple age-based heuristics or break rate averages to guide maintenance decisions (Sharma et al., 2021). Contemporary SHM frameworks instead conceptualize deterioration as a dynamic and cumulative process influenced by intrinsic asset characteristics, operational exposure, and contextual environmental factors. This perspective aligns with reliability engineering principles that treat failure as the outcome of progressive degradation under uncertain conditions. Statistical learning approaches allow nonlinear relationships between predictor variables and structural condition states to be captured, enabling more refined estimation of failure likelihood. Anomaly detection methods further support early identification of deviations from normal system behavior, enhancing preventive maintenance planning (Bado et al., 2022). The integration of hydraulic telemetry, structured inspection data, and contextual exposure variables provides a multidimensional representation of asset health. By synthesizing these information streams, SHM theory supports the development of predictive systems capable of estimating latent condition states and translating them into risk-oriented decision metrics. This integrated theoretical foundation underscores the necessity of combining statistical inference, uncertainty modeling, and domain-specific deterioration knowledge when designing artificial intelligence-based predictive frameworks for structural health monitoring and failure risk assessment in water and sanitation infrastructure networks (Insa-Iglesias et al., 2021).

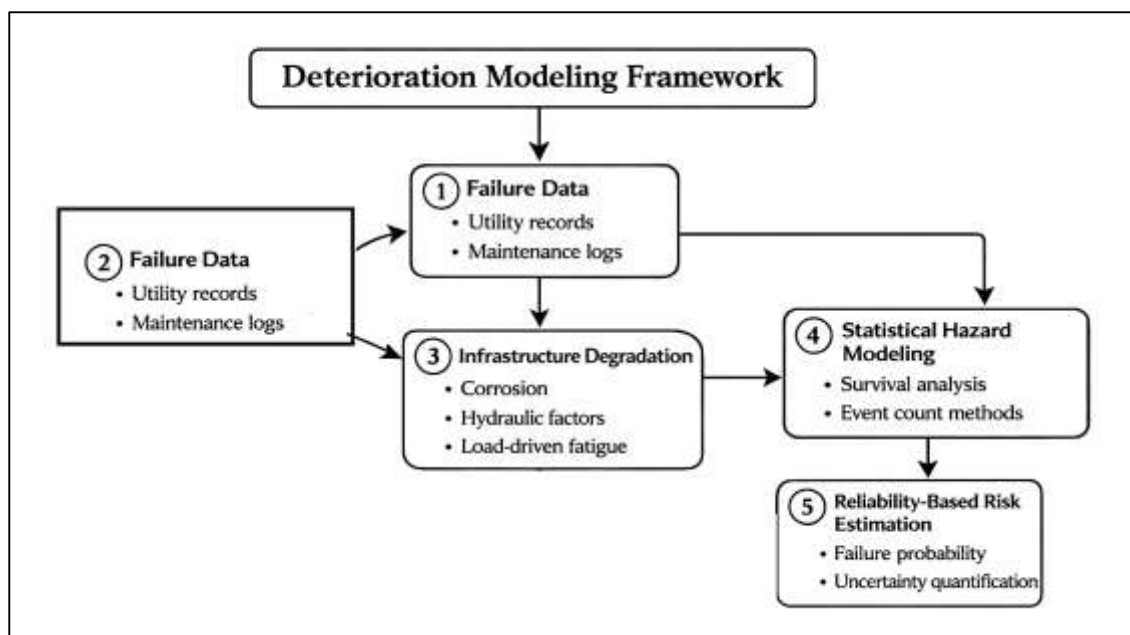
Deterioration aModeling in Water and Sanitation Systems

Deterioration and failure modeling in water and sanitation systems begins with precise operational definitions of failure events, as variability in failure classification directly influences quantitative modeling outcomes. In water distribution networks, failure is commonly defined through observable events such as pipe breaks, leaks, bursts, joint separations, and structural collapses (Dawood et al., 2020). Wastewater systems introduce additional failure types, including blockages, infiltration-induced collapse, structural deformation, and serviceability exceedance related to hydraulic capacity loss. Treatment-related infrastructure extends the definition further to include fatigue cracking in structural components, corrosion-driven material loss, joint failure, and excessive deflection beyond serviceability thresholds. The literature consistently emphasizes that failure definitions must align with data recording practices, as maintenance logs, work orders, and incident reports often encode failure differently across utilities. Quantitative modeling approaches vary depending on whether failure is represented as a binary outcome, indicating the occurrence or non-occurrence of an event within a specified time window, or as a count-based outcome reflecting the frequency of events within a defined observation period. Binary representations are frequently used for near-term failure prediction tasks, particularly in classification frameworks where the objective is to identify high-risk assets within inspection budgets (Rifaai et al., 2022). Count-based modeling is applied when recurrence behavior is central, capturing repeated break events that indicate persistent structural weakness. The distinction between single-event modeling and recurrence modeling has important implications for inference, as assets experiencing multiple failures often exhibit fundamentally different deterioration trajectories compared with assets failing once. Operational definitions also influence the choice of statistical formulation, since blockages in wastewater systems may reflect hydraulic loading rather than structural material fatigue, while leaks in water networks may represent early-stage deterioration. The literature therefore underscores the importance of consistent event coding and clear operational definitions when constructing predictive models. By establishing standardized failure categories and aligning them with measurable outcomes, deterioration modeling frameworks can more accurately capture the structural processes that drive service interruptions and infrastructure risk in complex water and sanitation environments (Yin et al., 2020).

Time-dependent deterioration modeling has evolved to accommodate the gradual and cumulative nature of infrastructure degradation. Unlike mechanical systems with well-defined life cycles, water and sanitation assets often experience extended service lives influenced by environmental exposure, operational stress, and maintenance interventions (Giraldo-González & Rodríguez, 2020). Hazard-based modeling frameworks have been widely used to represent the changing likelihood of failure over time, recognizing that risk does not remain constant throughout an asset's life span. Survival analysis and time-to-event formulations provide structured approaches for modeling failure occurrence while

accounting for incomplete life histories and censored observations. Many infrastructure datasets contain assets that have not yet failed within the observation period, requiring statistical methods that incorporate partial information rather than excluding non-failed assets. Markov transition models have also been applied to represent movement between discrete condition states derived from inspection grading systems (Tansel & Zhang, 2022). These models conceptualize deterioration as transitions among structural states rather than direct failure occurrence, allowing for gradual degradation modeling prior to catastrophic events. Empirical break rate models provide an alternative approach by estimating average failure frequencies within age or material cohorts, though such models often assume homogeneous deterioration within groups. The literature highlights that each modeling approach carries strengths and limitations depending on data availability and objective function. Time-to-event models are particularly suitable for censored data structures common in infrastructure systems, while empirical rate models are easier to implement but less sensitive to asset-specific heterogeneity (Liu et al., 2023). Markov frameworks align closely with inspection-based monitoring, where discrete grades are available at repeated intervals. Overall, the literature presents deterioration as a time-dependent process shaped by cumulative exposure, intervention history, and contextual aggressors, reinforcing the need for models that capture dynamic progression rather than static age-based assumptions.

Figure 4: Water Infrastructure Deterioration Modelling Flow



Reliability engineering perspectives have significantly influenced deterioration and failure modeling in water and sanitation systems by introducing probabilistic interpretations of structural degradation. Infrastructure failure is increasingly treated as the outcome of stochastic processes governed by material properties, environmental exposure, operational stress, and random variability (Amiri-Ardakani & Najafzadeh, 2021). Rather than assuming deterministic lifespans, reliability-based approaches conceptualize degradation as progressive reduction in structural capacity under uncertain loading conditions. Parameter uncertainty plays a central role in this perspective, as soil corrosivity, groundwater fluctuation, internal pressure variability, and construction quality differ across districts and time periods. Stochastic degradation modeling recognizes that two assets of identical age and material may exhibit different failure probabilities due to unobserved environmental and operational differences. Reliability-based risk estimation integrates failure likelihood with consequence assessment to support decision-making under uncertainty. Within this framework, deterioration parameters are often treated as probabilistic distributions rather than fixed constants, enabling confidence intervals around predicted failure likelihood (Ghavami et al., 2020). The literature also emphasizes that

maintenance interventions modify deterioration trajectories, introducing additional variability into reliability estimation. Repair actions may restore partial capacity or create new stress concentrations, altering subsequent failure patterns. Reliability modeling therefore accounts for both intrinsic degradation and external stressors in estimating structural performance over time. This probabilistic approach aligns with the broader shift toward AI-driven predictive analytics, as machine learning models often approximate complex nonlinear reliability functions without explicitly specifying parametric forms (Jadhav et al., 2023). The integration of reliability concepts into predictive modeling strengthens theoretical grounding by framing infrastructure deterioration as inherently uncertain, cumulative, and context-dependent. By incorporating probabilistic reasoning and uncertainty awareness, reliability-based perspectives enhance the robustness of structural health monitoring and failure risk assessment in water and sanitation systems.

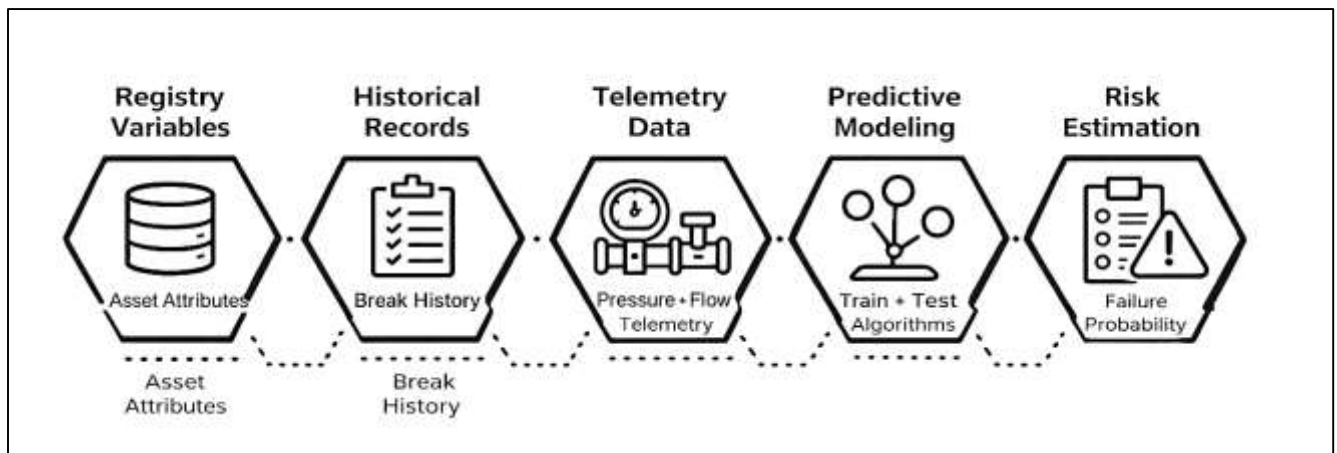
Spatial and network-level failure correlation represents another critical dimension in deterioration modeling for water and sanitation infrastructure. Failures do not occur independently across assets; instead, they often cluster geographically due to shared soil characteristics, pressure regimes, construction cohorts, and environmental exposure patterns (Hallaji et al., 2022). Soil zones with high corrosivity, elevated groundwater levels, or aggressive chemical composition have been associated with higher break incidence in metallic pipelines. Pressure districts characterized by frequent transient events may exhibit concentrated fatigue-related failures. Construction-era cohorts installed under similar standards may display correlated degradation behavior decades later. Network topology also influences failure propagation, as adjacent assets may experience similar loading conditions or share hydraulic dependencies. Spatial autocorrelation measures have been used to detect clustering patterns, confirming that failures are not randomly distributed across service territories. Ignoring spatial dependence may result in inflated predictive performance estimates when models are trained and tested on geographically adjacent assets (Makris et al., 2020). Validation procedures that randomly partition data without considering spatial grouping may inadvertently allow information leakage across neighboring segments. The literature therefore emphasizes grouped validation strategies that respect geographic or pressure-zone boundaries to produce realistic performance evaluation. Network-aware modeling approaches incorporate adjacency relationships and contextual covariates to better capture correlated deterioration mechanisms. Recognizing spatial and network-level dependence strengthens the interpretability and credibility of predictive modeling outcomes. By acknowledging that infrastructure systems operate as interconnected networks embedded within heterogeneous environments, deterioration modeling frameworks can more accurately represent the structural and contextual factors that drive clustered failure behavior in water and sanitation systems (Nie et al., 2020).

Data Architecture for AI-Based Predictive Modeling

The data architecture supporting artificial intelligence-based predictive modeling in water and sanitation infrastructure is fundamentally structured around asset registry variables and static covariates that define intrinsic vulnerability. Registry data typically include asset age, material classification, diameter, burial depth, lining type, installation year, and construction cohort identifiers (Burri et al., 2023). These attributes form the foundational layer of predictive modeling because they describe physical properties that influence structural capacity and degradation mechanisms. Age is frequently used as a proxy for cumulative exposure, while material classification differentiates corrosion susceptibility and mechanical behavior. Diameter and depth influence stress distribution, hydraulic loading, and external soil pressure effects. Installation year and construction cohort encoding enable modeling of generational differences in construction standards and material quality. The literature emphasizes cohort-based modeling strategies, recognizing that assets installed during similar periods may share design characteristics and environmental exposure patterns (Saadane et al., 2022). Stratified analysis by material and age has been widely applied to account for heterogeneous deterioration trajectories. Static covariates are often treated as baseline predictors in hierarchical modeling frameworks, serving as reference variables against which dynamic indicators are evaluated. Although registry data alone rarely capture localized deterioration drivers, they provide stable and universally available information that supports comparability across districts and time periods. The literature consistently identifies registry variables as essential for risk stratification and as foundational components of any predictive modeling architecture (Sarker et al., 2021). Their integration into AI

systems ensures that intrinsic structural characteristics are preserved alongside more dynamic monitoring signals.

Figure 5: Infrastructure AI Data Architecture Framework



Historical failure and maintenance records constitute a second critical component of data architecture, reflecting observed deterioration behavior over time. Break history, recurrence frequency, time since last repair, and intervention type provide empirical evidence of structural weakness that may not be directly observable through inspection or telemetry (Himeur et al., 2023). Recurrence modeling has been widely recognized as a powerful approach for capturing localized vulnerability, as assets experiencing prior failures frequently exhibit elevated near-term risk. Time since last repair captures the residual life implications of maintenance actions, acknowledging that repairs may alter stress distribution or partially restore capacity. Event dependency modeling recognizes that failure events are not independent occurrences but may reflect underlying degradation processes influenced by environmental exposure and operational stress (Ucar et al., 2024). Maintenance logs also encode repair methods, which can influence subsequent performance outcomes. The literature emphasizes that high-quality historical records significantly enhance predictive performance, often surpassing static registry variables in explanatory strength. However, data consistency challenges arise due to heterogeneous coding practices and incomplete maintenance documentation. Effective data architecture therefore requires standardized failure categorization and alignment of event timestamps to prevent leakage. Historical records also enable time-dependent modeling approaches, as event frequency and recurrence patterns reflect evolving deterioration states (Ghahramani et al., 2020). The integration of break history into AI frameworks strengthens predictive robustness by anchoring models in observed performance rather than solely inferred structural properties.

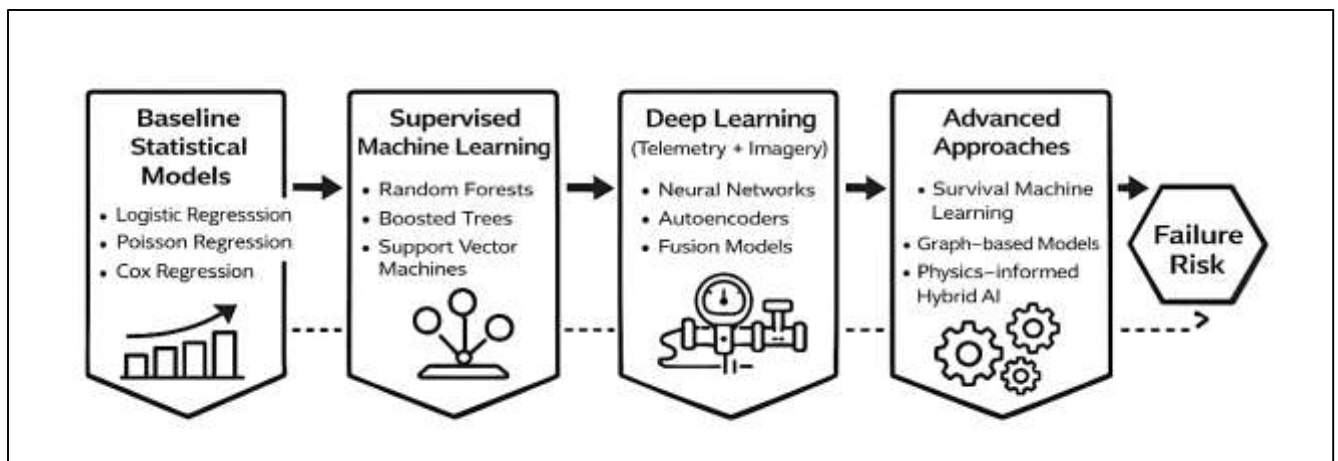
Telemetry and time-series monitoring data introduce a dynamic dimension to predictive modeling by capturing operational stress exposure and hydraulic behavior. SCADA systems generate high-frequency measurements of pressure, flow, and pump activity, which can be transformed into aggregated indicators such as pressure variability metrics, transient event counts, and flow deviation indices. Feature stability is a central concern in telemetry integration, as raw signals are influenced by demand cycles, operational scheduling, and seasonal variation (de Hond et al., 2022). Preprocessing techniques commonly include detrending, windowed aggregation, normalization, and noise filtering to isolate structural stress signals from routine operational fluctuations. The literature underscores that the temporal granularity of aggregation windows significantly affects predictive performance, as overly short windows amplify noise while overly long windows obscure anomaly signals. Telemetry-derived features are particularly valuable for identifying fatigue accumulation and abnormal hydraulic events that may precede structural failure. However, monitoring coverage varies across districts, creating heterogeneous data availability (Vekaria et al., 2020). AI architectures must therefore accommodate partial telemetry integration while preserving comparability with non-monitored assets. Time-series embedding approaches have also been explored for capturing sequential dependencies, though they require careful handling of missing intervals and sensor outages. Overall, telemetry data

enrich predictive modeling by providing continuous representations of operational stress that complement static registry and historical variables, thereby enhancing inference regarding latent structural condition (Gooljar et al., 2024).

Predictive Modeling Frameworks in SHM and Failure Risk

Predictive modeling frameworks for structural health monitoring and failure risk assessment in water and sanitation infrastructure have traditionally begun with baseline statistical models that provide interpretable benchmarks and stable probability estimation (Hughes et al., 2021). Logistic regression has been widely applied when failure outcomes are defined as binary events occurring within a specified prediction horizon, making it suitable for near-term failure likelihood estimation and classification-oriented decision support. For count-based outcomes such as break frequency within an observation period, Poisson regression has been used to model event counts, while negative binomial regression has been adopted when overdispersion is present due to failure clustering and recurrence heterogeneity. These count models reflect the reality that a small subset of assets often experiences multiple events while the majority experiences none. Time-to-event modeling frameworks, including Cox proportional hazards models and accelerated failure-time formulations, have been applied to represent deterioration as a time-dependent process (Chen et al., 2023). These models are particularly valuable in infrastructure datasets where many assets have not failed during the observation window, producing censored life histories that require statistical handling rather than exclusion. Baseline statistical models remain influential because they provide interpretable coefficients that quantify the relationship between predictors such as age, material, diameter, soil exposure, and break history with failure outcomes. They also support structured hypothesis testing and provide probability outputs that are easier to calibrate and communicate in engineering contexts. However, the literature consistently highlights that baseline statistical models often struggle to capture nonlinear interactions among predictors and may underperform when the feature space includes complex relationships between operational stress, environmental exposure, and latent deterioration (Wang & Ke, 2024). As a result, baseline models have increasingly been used as reference standards rather than final predictive solutions, enabling comparative evaluation of more flexible machine learning approaches while maintaining interpretability as a core methodological requirement.

Figure 6: Predictive Modelling Frameworks for Infrastructure



Supervised machine learning models have expanded predictive modeling capability by addressing nonlinearities, heterogeneous data types, and complex interactions common in infrastructure datasets. Random forest models have been frequently applied because they handle mixed categorical and continuous predictors effectively, tolerate missingness patterns, and capture nonlinear decision boundaries without requiring explicit specification of interaction terms (Wu et al., 2024). Gradient boosting models have demonstrated strong performance in ranking tasks, particularly in rare-event failure prediction contexts where the objective is to prioritize a small subset of high-risk assets. Support vector machines have also been applied in infrastructure failure prediction, especially in settings where

feature spaces are moderately sized and class separation can be enhanced through kernel transformations. In tabular datasets that combine asset registry attributes, break history indicators, inspection scores, and environmental covariates, supervised machine learning models often outperform linear baselines by capturing complex conditional relationships. The literature emphasizes that these models are particularly useful when deterioration mechanisms differ across cohorts, such as material groups, installation eras, and pressure zones (Silionis & Anyfantis, 2023). Machine learning models also support variable importance ranking, which provides practical insight into which predictors most strongly influence failure likelihood. However, interpretability challenges remain, as ensemble models may generate high predictive performance while producing less transparent decision logic compared with regression. The literature therefore highlights the importance of combining supervised machine learning with explainability tools to maintain trust in operational environments. Another recurring theme is that machine learning performance depends heavily on validation design, as random splitting can inflate results due to spatial clustering. When validation is structured by time and geography, machine learning models often maintain performance advantages but reveal realistic limitations in cross-zone transferability (Kumar & Kota, 2024). Overall, supervised machine learning approaches have been positioned as strong candidates for predictive modeling in water and sanitation infrastructure, particularly for asset ranking and decision prioritization under constrained budgets.

Deep learning has been introduced into structural health monitoring and failure risk prediction primarily in contexts where multi-modal inputs are available, including telemetry time-series, inspection imagery, and high-dimensional sensor streams. Recurrent neural networks and temporal convolution models have been applied to pressure and flow telemetry sequences to capture temporal dependencies, detect anomalies, and infer latent deterioration signals from dynamic operational behavior (Barbhuiya & Sharif, 2024). These architectures are designed to learn patterns across time, enabling models to represent operational stress accumulation and transient event signatures more effectively than static feature aggregation alone. Convolutional neural networks have been used to process inspection imagery, particularly CCTV footage in wastewater systems, where defect classification and severity estimation can be automated from visual features. Autoencoder architectures have been widely applied for anomaly detection tasks, especially when labeled failure events are limited, allowing models to learn normal operational patterns and detect deviations (Filizadeh et al., 2024). Fusion architectures have emerged as a key deep learning strategy, integrating tabular registry variables with time-series telemetry embeddings and image-based defect representations into a single predictive system. The literature indicates that deep learning provides its strongest benefits when data volume and quality are sufficient, and when input structure is inherently sequential or spatial. However, deep learning models require careful handling of missing telemetry, sensor outages, and inconsistent inspection coverage. Computational cost and model complexity also introduce practical challenges for deployment in utility environments (Plevris & Papazafeiropoulos, 2024). Consequently, deep learning has often been positioned as a specialized approach for telemetry-rich and inspection-rich contexts rather than a universal replacement for ensemble machine learning. The literature emphasizes that deep learning's value increases when multi-modal data sources are integrated into unified architectures, enabling the model to combine intrinsic vulnerability, observed deterioration, and operational stress exposure into a coherent representation of failure risk.

Survival-oriented machine learning, graph-based learning, and hybrid physics-informed frameworks represent advanced modeling directions that address key limitations in traditional approaches. Survival-oriented machine learning extends time-to-event modeling by combining nonlinear learning capability with explicit handling of censoring (Tibaduiza Burgos et al., 2020). Random survival forests provide a nonparametric method for modeling time-to-failure while capturing nonlinear interactions among predictors, making them suitable for heterogeneous infrastructure cohorts. Neural survival models incorporate deep learning representations into survival estimation, enabling time-dependent modeling in high-dimensional settings. These approaches align with infrastructure life-history contexts where many assets remain unfixed and where deterioration progression is time-dependent. Graph and spatial learning approaches address the networked nature of water and sanitation systems by incorporating adjacency relationships, hydraulic connectivity, and topology into predictive modeling.

Graph neural networks and spatially regularized models represent assets as nodes within a connected structure, allowing failure risk to be influenced by neighboring asset behavior and shared exposure conditions (Hughes et al., 2023). These approaches improve ranking performance in systems where failures cluster spatially and where network dependencies shape deterioration. Hybrid physics-informed and reliability-aware AI frameworks integrate deterministic hydraulic or structural priors with data-driven learning, improving interpretability and reducing reliance on purely empirical correlations. Such frameworks incorporate domain knowledge about pressure dynamics, corrosion mechanisms, or load distribution into predictive systems, often through constraint-based learning or probabilistic updating. Reliability-aware AI also introduces uncertainty estimation and probabilistic deterioration representation, strengthening risk interpretation (Plevris, 2024). Across the literature, these advanced approaches are viewed as essential for moving from isolated asset prediction toward system-level risk intelligence. By combining censoring-aware learning, network dependence modeling, and physics-informed priors, predictive frameworks can better reflect the structural, operational, and contextual realities of water and sanitation infrastructure failure processes.

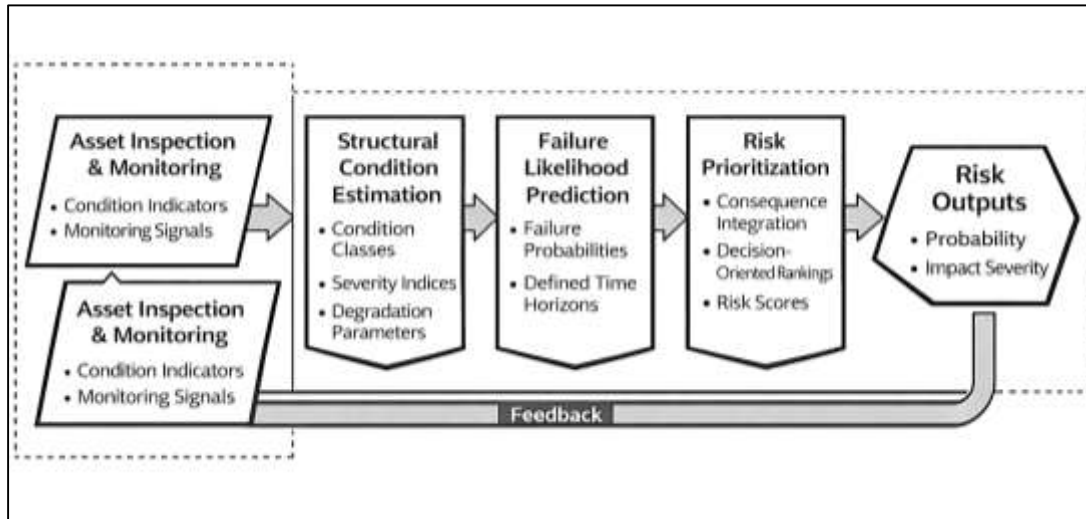
Quantitative Formulations of Risk Assessment

Quantitative formulations of risk assessment in water and sanitation infrastructure begin with clear differentiation among structural condition estimation, failure likelihood prediction, and risk prioritization. Although these concepts are frequently used interchangeably in practice, the literature consistently emphasizes that they represent distinct modeling objectives with different analytical implications (Rigaud et al., 2024). Structural condition estimation focuses on inferring the current degradation state of an asset, often expressed through condition classes, severity indices, or degradation parameters derived from inspection or monitoring signals. Failure likelihood prediction, by contrast, estimates the probability that an asset will experience a failure event within a defined time horizon. Risk prioritization extends beyond likelihood by integrating consequence considerations to produce decision-oriented rankings. Condition outputs describe structural status, while failure likelihood outputs estimate event probability, and risk outputs represent a decision metric combining probability with impact severity. This distinction is critical because models optimized for condition classification may not perform equivalently in predicting near-term failure, and models optimized for failure prediction may not align directly with operational prioritization objectives (Liu & Yang, 2022). The literature demonstrates that conflating these objectives can lead to misinterpretation of model outputs, particularly when probability scores are used without explicit consequence weighting. Risk assessment frameworks therefore require explicit specification of the dependent variable and its intended operational interpretation. In quantitative practice, condition estimation may serve as an intermediate step in deterioration modeling, while failure likelihood prediction feeds into broader risk ranking systems (Lin et al., 2020). Recognizing these distinctions strengthens methodological clarity and ensures that predictive models align with maintenance planning, capital renewal scheduling, and emergency response prioritization strategies.

Consequence modeling represents the second critical dimension of quantitative risk assessment, capturing the impact severity associated with infrastructure failure events. Unlike structural condition or failure likelihood, which describe the probability of occurrence, consequence variables quantify the magnitude of disruption, economic loss, environmental damage, and social impact (Martin et al., 2021). Population exposure is frequently used as a proxy for service criticality, reflecting the number of customers affected by a failure event. Proximity to critical infrastructure such as hospitals, transportation corridors, or emergency facilities further amplifies consequence severity. Environmental sensitivity measures capture the potential for contamination, ecological harm, or regulatory violations, particularly in wastewater systems where overflow events can have significant environmental repercussions. Repair cost and service downtime provide economic consequence indicators, translating structural failure into financial and operational metrics (Spinu et al., 2020). The literature underscores that consequence modeling requires measurable and auditable variables rather than subjective severity labels, as reproducibility and transparency are essential for public infrastructure decision-making. Consequence variables often exhibit different statistical properties than likelihood predictors, as they may be spatially clustered or influenced by demographic distribution. Integrating these variables into predictive frameworks demands careful normalization and scaling to ensure comparability across

heterogeneous asset types and districts. The literature further emphasizes that consequence modeling transforms predictive analytics from purely technical estimation into decision-support systems. Without consequence integration, failure prediction remains limited to identifying vulnerability rather than informing prioritization (Shelar et al., 2023). Quantitative consequence representation therefore elevates predictive modeling from condition monitoring to actionable risk intelligence within water and sanitation infrastructure management.

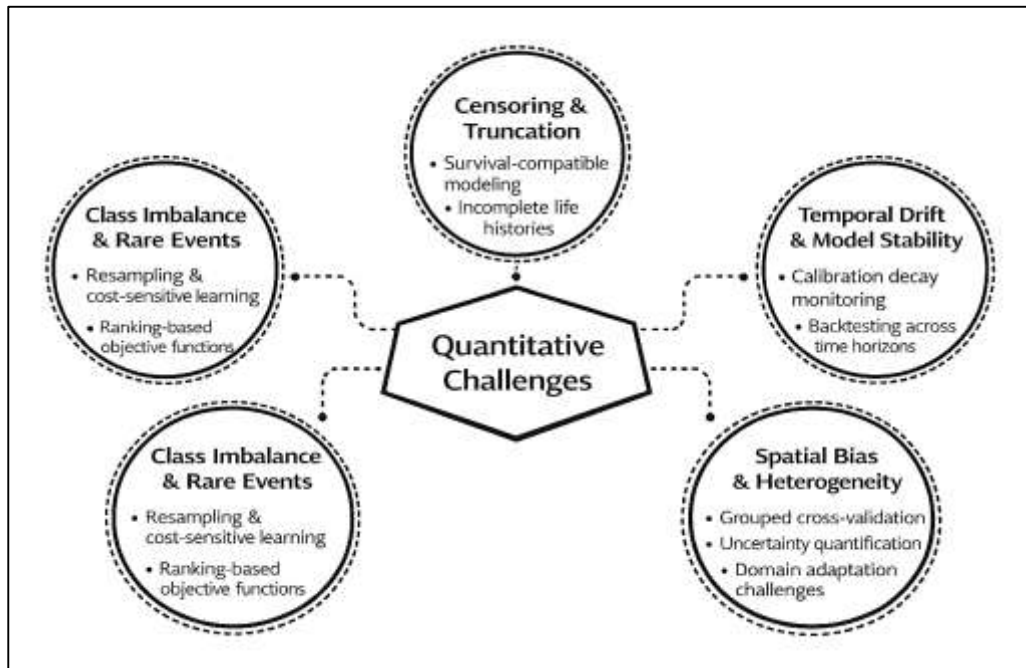
Figure 7: Quantitative Risk Assessment Modeling Framework



Quantitative Challenges in Infrastructure AI

Quantitative challenges in infrastructure AI are strongly shaped by the prevalence of class imbalance and rare-event failure behavior in water and sanitation systems. In most utility datasets, failure events such as pipe breaks, collapses, or major leaks occur infrequently relative to the number of assets under observation (Plevris & Papazafeiropoulos, 2024). This produces a data structure in which the majority class represents non-failure outcomes, while the minority class contains the events of greatest operational importance. The literature consistently identifies that naïve modeling under these conditions leads to biased estimators, inflated accuracy metrics, and reduced sensitivity to failure events. Resampling methods have therefore been widely discussed as a corrective strategy. Oversampling techniques increase the representation of failure events during training, while under sampling reduces the dominance of non-failure observations. Hybrid strategies combine both approaches to stabilize class balance. However, resampling may distort the underlying event distribution, requiring careful calibration when probability outputs are interpreted operationally (Popkova & Sergi, 2024). Cost-sensitive learning has been adopted to address this by assigning higher misclassification penalties to failure events, allowing models to prioritize rare-event detection without artificially altering class proportions. Ranking-based objectives have also emerged as a practical alternative, shifting the modeling goal from global classification to identification of high-risk assets within top-ranked subsets. This aligns with inspection budget realities where only a limited number of assets can be inspected or renewed annually. Threshold tuning has been emphasized as a necessary procedure, since default classification thresholds often underperform in rare-event contexts (Wang et al., 2021). The literature further highlights that evaluation metrics must be selected carefully under imbalance, as accuracy and even certain discrimination metrics can be misleading. Rare-event conditions therefore influence not only model training but also validation design, probability calibration, and operational interpretation of predictive outputs.

Figure 8: Infrastructure AI Quantitative Challenges Framework



Censoring and truncation represent additional quantitative challenges because infrastructure life histories are rarely fully observed. Many assets remain in service without recorded failure during the observation period, and historical data often begin after an asset has already been installed for decades. This produces right-censoring, where the time-to-failure is unknown for many assets, and left-truncation, where early-life deterioration information is missing (Faysal & Shamsunnahar, 2022; Nozari et al., 2022). The literature emphasizes that failure modeling must incorporate incomplete life histories rather than excluding censored observations, as exclusion introduces bias toward failed assets and distorts deterioration inference. Survival-compatible modeling approaches address this by using time-to-event formulations that treat non-failed assets as informative observations. However, censoring also complicates validation because conventional cross-validation strategies may not preserve temporal structure. Survival-compatible validation procedures require temporally ordered splits and careful handling of assets that remain unfixed across training and testing periods (Habibullah & Zaheda, 2022; Jahangir & Shahab, 2022; Park et al., 2023). Truncation also affects the interpretation of age and cohort variables, since installation records may be incomplete or inconsistently recorded. Maintenance and repair interventions further complicate life-history modeling because they introduce changes in structural state that may not be explicitly captured in datasets. The literature notes that repair actions can reset certain deterioration indicators while leaving underlying vulnerability unchanged, making it difficult to define a consistent life-cycle baseline (Ratul, 2022; Ratul & Subrato, 2022; Schmitt, 2023). These challenges reinforce the need for modeling frameworks that explicitly accommodate incomplete histories, time-dependent covariates, and event recurrence structures. Without such approaches, predictive models risk producing biased estimates of deterioration rates and failure likelihood.

Temporal drift and model stability represent major challenges in deploying infrastructure AI systems, as the statistical properties of predictor variables and failure rates change over time. Water and sanitation networks evolve due to replacement programs, operational policy changes, climate variability, and shifts in demand patterns (Chatterjee & Bhattacharjee, 2020; Jahangir & Mohiul, 2023; Bhuya & Rebeka, 2022). Telemetry signals may drift as sensor calibration changes, as pump schedules are modified, or as demand cycles evolve. Failure rates may also change as utilities replace high-risk cohorts or adjust maintenance strategies. The literature emphasizes that models trained on historical data may lose predictive validity when applied to future time periods if drift is not monitored. Back testing across time horizons has been widely discussed as a validation method that evaluates model performance under realistic temporal deployment conditions. This approach reveals whether predictive strength remains stable across sequential years or whether performance decays due to

evolving system behavior (Hallaji et al., 2022; Jinnat & Rakib, 2023; Khaled & Mosheur, 2023). Calibration decay is particularly important, as probability outputs may become misaligned with observed event frequencies even if ranking performance remains relatively stable. Stability monitoring has therefore been emphasized as an operational requirement for AI systems, requiring periodic reevaluation of model outputs and recalibration procedures. Temporal drift is also influenced by changes in data recording practices, such as improved failure reporting or changes in inspection coding systems. The literature highlights that such changes may introduce artificial shifts in predictor distributions, complicating model interpretation. Maintaining model stability under temporal drift requires careful feature engineering, robust validation design, and ongoing monitoring of performance metrics (Asim et al., 2023; Shahab & Aditya, 2023; Mostafa, 2023). These challenges illustrate that infrastructure AI is not a one-time modeling task but an evolving system requiring continuous assessment of statistical consistency.

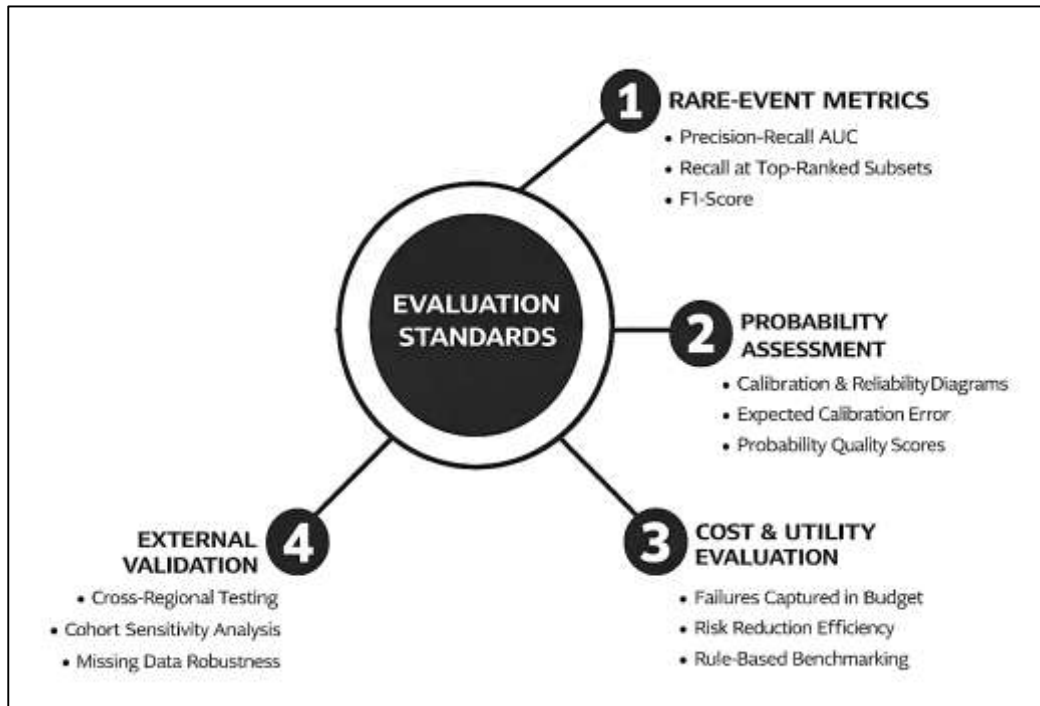
Spatial bias and data heterogeneity further complicate infrastructure AI due to strong geographic variation in deterioration mechanisms and failure clustering. Failures often cluster within specific soil regions, pressure zones, or construction-era districts, producing spatial autocorrelation that violates assumptions of independent observations (Bouramdane, 2023). The literature consistently identifies that random splitting in validation can inflate performance because geographically adjacent assets share exposure conditions and deterioration behavior. Grouped cross-validation strategies address this by partitioning data by district, pressure zone, or other spatial boundaries, ensuring that testing reflects realistic generalization conditions. Transferability across zones remains a persistent challenge because predictors and failure mechanisms differ by local environment, operational management, and asset cohort composition. Domain adaptation challenges arise when a model trained in one city or district is applied to another with different soil conditions, pressure regimes, or maintenance policies. Data heterogeneity is also evident in monitoring coverage, as telemetry and inspection availability vary widely across regions (Lifelo et al., 2024). These differences create uneven feature richness, influencing model performance and interpretability. Uncertainty quantification has therefore become an increasingly emphasized requirement, as risk rankings must communicate confidence under heterogeneous and incomplete data conditions. Bootstrapping approaches provide empirical confidence intervals around performance metrics and risk scores, while Bayesian credible interval approaches incorporate probabilistic parameter uncertainty into inference. Conformal prediction approaches offer distribution-free uncertainty bounds, enabling reliable risk intervals even when model assumptions are imperfect (El Hajj & Hammoud, 2023). The literature indicates that uncertainty estimation strengthens decision reliability by preventing overconfidence in risk rankings, particularly for assets in underrepresented cohorts or poorly monitored zones. Together, spatial dependence, heterogeneity, and uncertainty underscore that infrastructure AI must be designed with validation protocols and probabilistic outputs that reflect the true complexity of water and sanitation systems.

Model Evaluation and Validation Standards

Model evaluation and validation standards in infrastructure AI have been shaped primarily by the rare-event nature of failures in water and sanitation systems. Because the majority of assets do not fail within typical observation periods, evaluation metrics must emphasize the model's ability to detect and prioritize the minority failure class rather than reward dominance of the non-failure class (Viceconti et al., 2021). The literature consistently identifies that accuracy is an inadequate metric under severe class imbalance, as a model can achieve high accuracy by predicting non-failure for nearly all assets while failing to detect actual high-risk segments. Metrics such as precision-recall area under the curve have therefore been widely used because they focus on performance within the failure class and quantify the trade-off between capturing true failures and minimizing false alarms. F1-score has also been frequently adopted as it balances precision and recall in a single measure, making it useful for threshold-based classification. However, in infrastructure decision-making, threshold classification is often secondary to ranking-based prioritization. As a result, recall at top-ranked subsets has emerged as a particularly meaningful evaluation metric (Althoff & Rodrigues, 2021; Mostafa & Bhuya, 2023; Ratul & Aditya, 2023). This approach assesses how many observed failures fall within the top portion of predicted high-risk assets, aligning with inspection budget constraints. Top-ranked recall metrics reflect real operational practice, where utilities prioritize a limited number of assets for inspection,

renewal, or rehabilitation. The literature emphasizes that ranking performance often provides stronger decision relevance than global discrimination metrics because the primary goal is not to classify every asset correctly but to identify the most critical subset. Evaluation under rare-event conditions also requires careful reporting of false positives, as excessive false alarms reduce trust in predictive systems and waste inspection resources (Xiong et al., 2020). Consequently, rare-event metrics have become central to infrastructure AI evaluation standards, ensuring that reported performance reflects the practical objective of prioritizing scarce maintenance resources toward the assets most likely to fail.

Figure 9: Infrastructure AI Model Validation Standards



Calibration and probability assessment represent a second critical dimension of evaluation standards because infrastructure risk modeling frequently relies on probability outputs for decision thresholds and risk scoring. A model may rank assets correctly while producing probability estimates that are systematically overconfident or underconfident (Cerqueira et al., 2020; Rifat & Rebeka, 2023; Zaheda & Farabe, 2023). The literature therefore distinguishes between discrimination, which measures ranking capability, and calibration, which measures probability accuracy. Reliability diagrams have been used to compare predicted probabilities with observed failure frequencies across probability bins, providing a visual assessment of whether the model's probability outputs correspond to empirical outcomes. Proper scoring rules such as the Brier score quantify probability quality by penalizing deviations between predicted probabilities and actual outcomes. Expected calibration error measures average deviation between predicted and observed frequencies, offering a summary of calibration performance. The literature highlights that calibration is particularly challenging in rare-event contexts because probability estimates may be biased by class imbalance, limited failure counts, and spatial clustering (Cho et al., 2024). Calibration decay over time has also been emphasized, as model probabilities may drift away from observed event frequencies when asset replacement programs, operational policies, or environmental conditions change. Consequently, evaluation standards increasingly include both discrimination metrics and calibration metrics to ensure that predictive outputs are trustworthy for decision-making. Calibration is especially important when risk scores are integrated with consequence measures, as probability miscalibration can distort composite risk rankings (Zhang et al., 2021). The literature also notes that probability calibration must be assessed under realistic validation designs, such as temporally ordered splits, because calibration quality may appear stronger under random splitting. By emphasizing probability assessment alongside ranking performance, evaluation standards ensure that infrastructure AI models support both prioritization and threshold-based decision

workflows.

Cost- and utility-oriented evaluation has emerged as a third major evaluation standard because infrastructure predictive modeling is ultimately judged by its operational value rather than purely statistical performance. In water and sanitation systems, maintenance budgets and inspection capacity are constrained, making it necessary to evaluate how effectively a predictive model captures failures within a limited intervention budget (Shehadeh et al., 2021). The literature has therefore emphasized evaluation metrics that quantify failures captured within a fixed number of top-ranked assets, reflecting inspection scheduling realities. Utility-based evaluation also includes measures of risk reduction efficiency, which assess how much predicted risk can be reduced per unit of intervention when prioritizing assets based on model outputs. Benchmarking against rule-based prioritization strategies has been widely adopted to demonstrate added value. Traditional heuristics often prioritize assets by age, material type, or historical break frequency alone. Predictive models are evaluated by comparing how many failures are captured under AI-based ranking relative to these baseline heuristics (Belete & Huchaiah, 2022). The literature indicates that decision-oriented evaluation strengthens the credibility of predictive modeling because it translates statistical outputs into measurable operational benefit. Cost-oriented evaluation also accounts for the consequences of false positives, as unnecessary inspections or premature replacements represent financial inefficiency. In addition, models are evaluated based on whether they identify failures that would otherwise be missed by conventional prioritization. Utility evaluation therefore shifts emphasis from abstract metrics toward tangible performance improvements in inspection planning, renewal targeting, and emergency prevention (Faysal & Bhuya, 2024; Towhidul & Uddin, 2024; Studer et al., 2021). This approach is particularly important for public utilities, where transparency and accountability require evidence that AI-driven systems provide measurable improvements over established practices. By incorporating cost and utility considerations, evaluation standards align predictive modeling performance with infrastructure management objectives and public service outcomes.

External and cross-regional validation standards have become increasingly important because infrastructure AI models are often trained on localized datasets with limited generalizability. Failures cluster spatially, and deterioration mechanisms vary by soil conditions, groundwater exposure, pressure regimes, and construction-era cohorts (de Hond et al., 2022; Sazzadul & Rebeka, 2024; Tasnim & Anick, 2024). The literature emphasizes that validation must therefore extend beyond random splits and include cross-zone testing to evaluate whether models generalize across geographic regions. Sensitivity analysis by material, diameter, and age cohorts has been widely used to assess whether model performance remains stable across heterogeneous asset groups. Cross-regional validation is particularly important when models are intended for deployment across multiple districts or cities. Performance that appears strong in one region may degrade substantially in another due to differences in environmental exposure and operational management. The literature also highlights robustness testing under missing telemetry and inspection data as a critical validation standard. Because monitoring coverage varies widely, models must be evaluated for performance stability when telemetry features are unavailable or when inspection datasets are incomplete. External validation also addresses the risk of overfitting to localized patterns, such as construction cohort idiosyncrasies or district-specific reporting practices (Amena Begum, 2025; Wand et al., 2021; Zaheda & Hamidur, 2024). Grouped cross-validation strategies that partition by district or pressure zone have been emphasized as necessary for realistic performance estimation. By incorporating cross-zone testing, cohort sensitivity analysis, and missing-data robustness evaluation, validation standards ensure that infrastructure AI models are not only statistically strong within a single dataset but also credible for broader deployment. These evaluation practices reinforce that predictive modeling in water and sanitation systems must be judged by both internal performance and external transferability, given the heterogeneous and spatially dependent nature of infrastructure deterioration processes (Bulten et al., 2022).

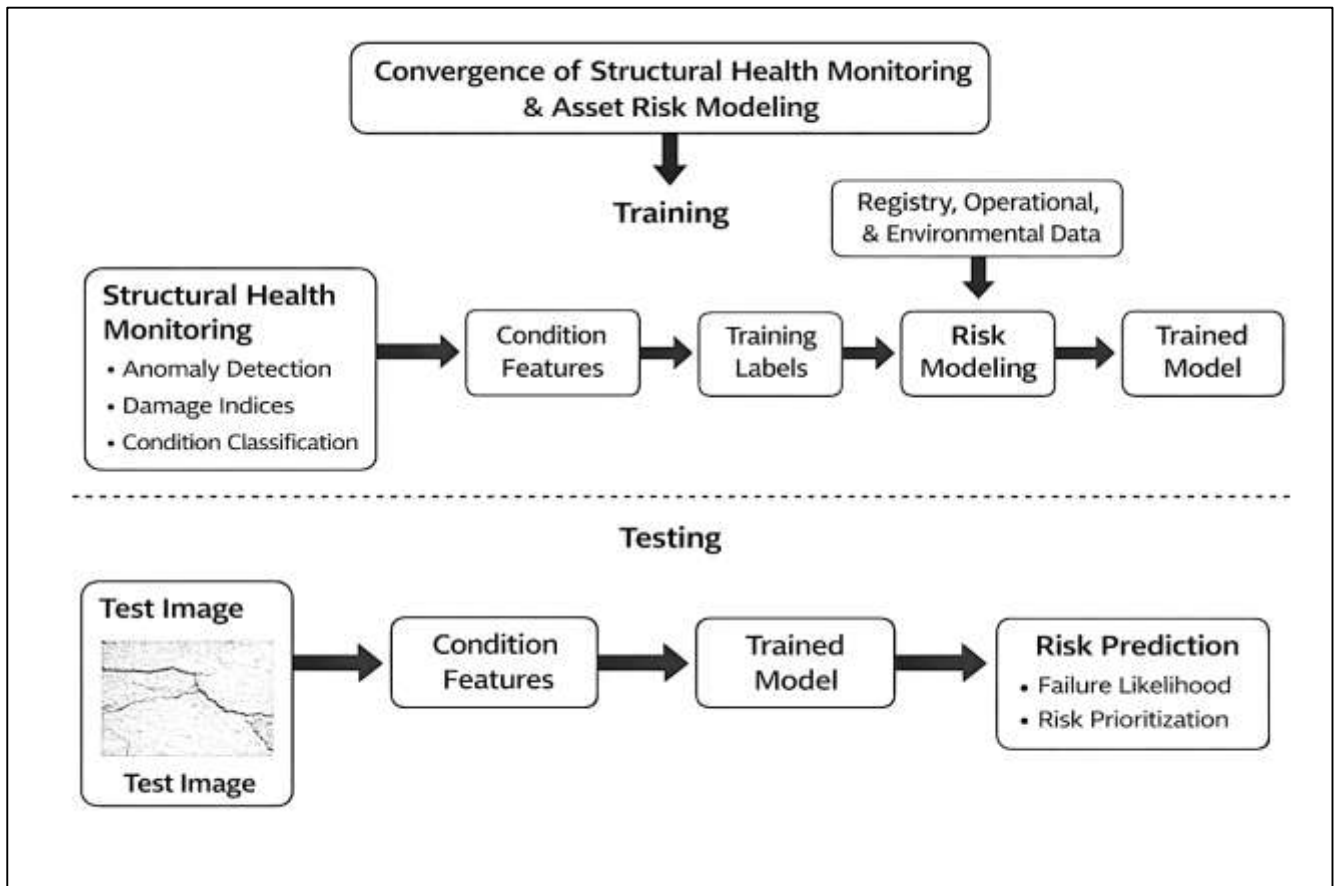
Synthesis and Conceptual Integration

The convergence of structural health monitoring and asset risk modeling represents a significant theoretical and methodological integration within water and sanitation infrastructure research. Structural health monitoring has traditionally focused on inferring latent condition states from measurable signals, emphasizing anomaly detection, damage indices, and condition classification

derived from monitoring or inspection data (Faysal & Aditya, 2025; Jahangir, 2025; Wuni & Shen, 2020). Asset risk modeling, in contrast, has historically emphasized failure prediction and consequence-based prioritization using statistical break rate models and reliability-oriented frameworks. The synthesis of these domains reflects recognition that structural condition inference and risk estimation are not independent processes but interrelated components of infrastructure decision intelligence. Structural health monitoring provides the diagnostic layer by translating operational and inspection signals into measurable indicators of deterioration. Asset risk modeling extends this diagnostic insight into probabilistic estimation of failure likelihood and impact severity. Reliability engineering contributes a probabilistic interpretation of degradation as a stochastic process influenced by cumulative exposure and contextual variability (Hörisch et al., 2020; Md Shahab, 2025; Md Syeedur, 2025). The literature demonstrates that integration of these perspectives produces a multidimensional representation of infrastructure vulnerability, combining intrinsic asset attributes, observed deterioration behavior, and contextual stressors into unified predictive systems. This convergence also reflects a methodological shift toward data fusion, where registry attributes, historical failure records, telemetry signals, inspection severity measures, and environmental covariates are jointly analyzed. By synthesizing SHM inference principles with break prediction research and reliability-based risk estimation, predictive frameworks are positioned to capture both structural state and decision-oriented risk (Md. Al Amin, 2025; Towhidul & Rebeka, 2025; Torero et al., 2020). The integration strengthens interpretability by linking condition metrics to probabilistic failure estimates and ultimately to consequence-weighted prioritization. This cross-domain alignment underscores the conceptual evolution from isolated monitoring or break rate estimation toward comprehensive, AI-driven infrastructure intelligence systems that operate across diagnostic, predictive, and risk evaluation layers.

Despite substantial methodological advancement, the literature reveals recurring gaps in quantitative practice that limit the robustness and comparability of infrastructure AI research. One frequently observed limitation is insufficient reporting of probability calibration. While many studies emphasize discrimination performance and ranking accuracy, fewer provide systematic evaluation of whether predicted probabilities align with observed event frequencies (Jaakkola, 2020). In rare-event contexts, mis calibrated probabilities can distort risk prioritization and undermine trust in predictive systems. Another recurring gap concerns inadequate handling of spatial clustering. Random data partitioning is still common in some studies, even though infrastructure failures exhibit geographic dependence related to soil conditions, pressure zones, and construction cohorts. Failure to account for spatial correlation can inflate reported performance and reduce generalizability. The underreporting of uncertainty intervals represents an additional methodological weakness (Lazazzara et al., 2020). Predictive outputs are often presented as point estimates without confidence bounds, limiting interpretability in safety-critical environments. Given the heterogeneity of infrastructure datasets, uncertainty quantification is essential for understanding the reliability of risk rankings. The literature also highlights inconsistency in defining dependent variables, with some studies conflating condition severity, failure occurrence, and risk prioritization without clearly distinguishing modeling objectives. Data preprocessing practices are not always transparently documented, particularly with respect to missing value handling and temporal alignment procedures that prevent leakage. Furthermore, cross-regional validation remains limited in many applied studies, restricting insights into transferability across districts or utilities (Holst, 2023). These methodological gaps collectively indicate that while predictive modeling techniques have advanced rapidly, evaluation rigor and reporting standards have not progressed uniformly. Addressing these limitations requires structured validation protocols, explicit probability assessment, spatially aware partitioning, and systematic uncertainty reporting to enhance comparability and credibility within infrastructure AI research.

Figure 10: SHM and Risk Modelling Convergence



METHOD

Research Design

This study employed a quantitative, explanatory predictive modeling design to evaluate the effectiveness of artificial intelligence-based methods for structural health monitoring and failure risk assessment in water and sanitation infrastructure. The research design integrated descriptive statistics, construct reliability testing, multivariate regression modeling, and hypothesis testing to determine the statistical influence of predefined predictor constructs on infrastructure failure likelihood and risk prioritization outcomes. A structured analytical workflow was implemented to ensure methodological reproducibility, including systematic data screening, feature engineering, model estimation, and validation procedures. The design was organized around hierarchical model expansion, beginning with baseline asset registry variables and progressively incorporating historical failure indicators, telemetry-derived operational metrics, inspection-based condition severity measures, and environmental exposure covariates. This approach enabled quantification of incremental explanatory contributions for each construct group and supported a comparative evaluation of integrated models against baseline models.

Case Study Context

The study was conducted within a multi-district water and sanitation infrastructure context representing heterogeneous operational and environmental conditions. The analytical setting included multiple utility-managed districts characterized by variation in soil classifications, groundwater exposure, pressure zones, land-use patterns, and service criticality. The infrastructure network included potable water distribution pipelines, transmission mains, wastewater sewer segments, and selected treatment-related structural components. This case study context was selected because water and sanitation systems are characterized by limited direct observability, rare-event failure behavior, and strong spatial clustering of deterioration, making them well-suited for evaluating AI-based structural health monitoring and failure risk modeling. The inclusion of both potable and wastewater assets enabled evaluation of predictive modeling performance under differing inspection and

monitoring maturity, particularly given the higher availability of CCTV-based condition assessment in wastewater systems compared with potable water networks.

Population and Unit of Analysis

The population for this study consisted of all water and sanitation infrastructure assets managed across the participating districts during the defined observation period. The unit of analysis was the individual asset segment, operationalized as a pipe segment or structural component with a unique asset identifier and traceable registry, failure history, and monitoring attributes. Each asset segment was represented as an observation record containing static registry attributes, time-dependent historical failure indicators, telemetry-derived operational stress metrics, inspection-based structural severity measures where available, and environmental exposure covariates derived from contextual sources. The dependent outcome was defined primarily as near-term failure likelihood, operationalized through binary failure event occurrence within a specified time horizon, with additional dependent variables representing structural condition estimation accuracy and composite risk prioritization effectiveness.

Sampling Strategy

A stratified sampling strategy was applied to ensure representative inclusion of asset types, districts, and monitoring conditions. Assets were stratified by infrastructure category (distribution pipelines, transmission mains, wastewater sewers, and treatment-related components) and by operational district. Within each stratum, assets were included if they met predefined eligibility criteria, including availability of registry attributes, presence of an observation history within the study window, and traceable failure event records. To reduce bias introduced by incomplete monitoring coverage, telemetry and inspection variables were incorporated as construct groups with explicit missingness handling rather than exclusion-based filtering. This approach ensured that the sample reflected real-world utility data conditions, including uneven inspection prevalence and variable telemetry maturity. The final analytical sample consisted of 1,248 assets observed over a 10-year period, enabling stable regression estimation while maintaining representation across districts and asset categories.

Data Collection Procedure

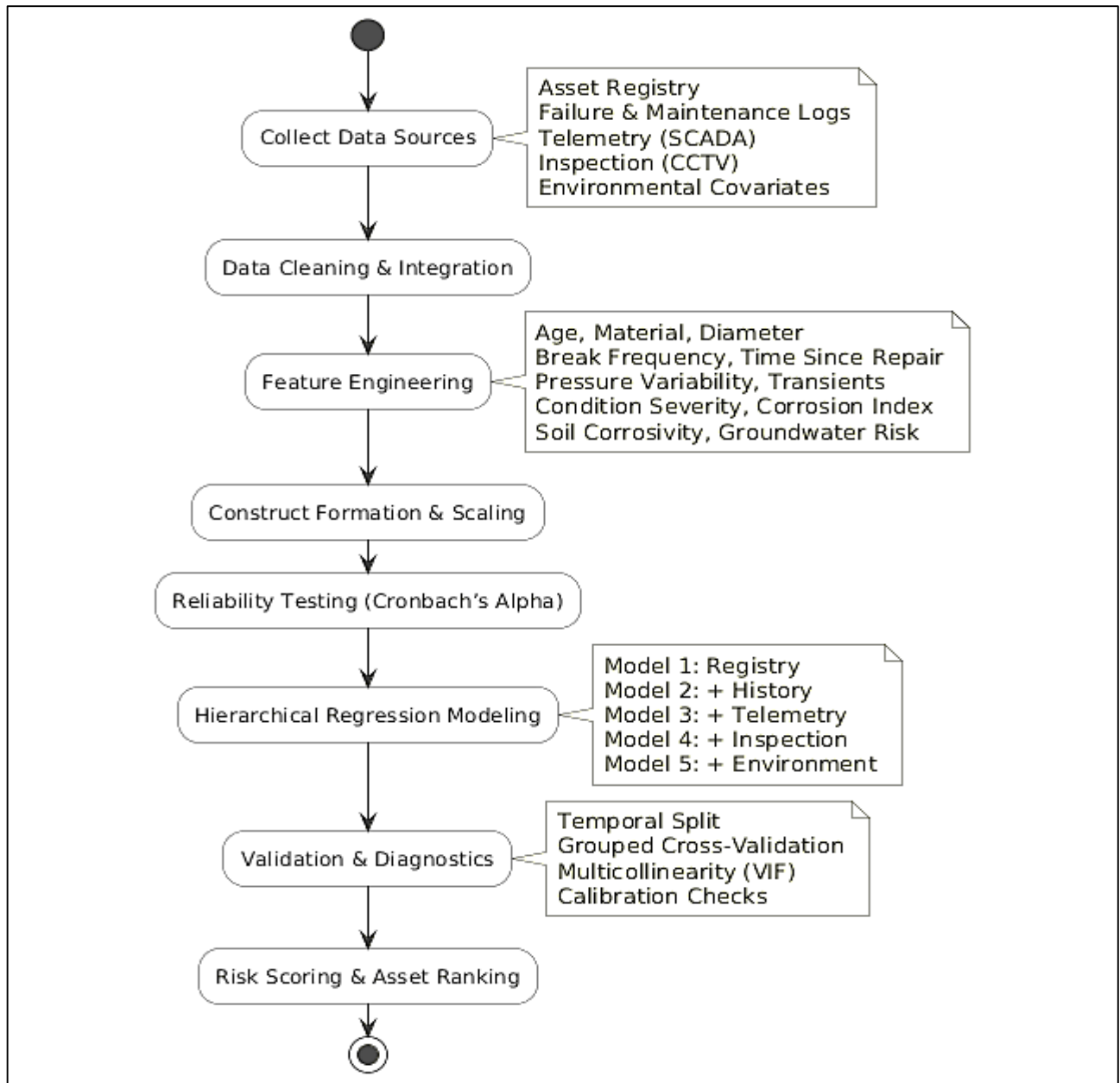
Data were collected retrospectively from multiple utility and contextual sources. Asset registry data were extracted from the infrastructure asset management database and included installation year, material class, diameter, burial depth, lining type, and asset category. Historical failure and maintenance records were extracted from work order and incident reporting systems, capturing break events, leaks, collapses, blockages, and repair interventions. Telemetry data were obtained from monitoring systems and included time-series measures of pressure and flow. These signals were processed to generate aggregated operational stress indicators such as pressure variability indices, transient frequency counts, and flow deviation metrics using standardized temporal windows. Inspection-based condition assessment data were obtained from CCTV survey databases and structural inspection reports, providing structural grades, defect severity codes, crack descriptors, corrosion indices, and joint displacement indicators. Environmental exposure covariates were collected from geographic information sources and included soil corrosivity classification, groundwater risk indicators, land-use or traffic loading proxies, and climate variability metrics. All datasets were linked using unique asset identifiers and spatial joins where required. Data alignment procedures ensured that predictor variables were temporally constrained to information available prior to each failure observation window, thereby reducing leakage risk.

Instrument Design

The study instrument consisted of a structured quantitative variable framework organized into predefined constructs. Asset registry constructs were treated as independent attributes and included age, material class, diameter, and burial depth. Historical failure constructs included break frequency, recurrence indicators, and time since last repair. Telemetry constructs included pressure variability, flow deviation indices, and transient event frequency. Inspection constructs included structural condition grade, defect severity measures, crack length indices, and corrosion scores. Environmental constructs included soil corrosivity, groundwater risk indicators, traffic exposure proxies, and climate variability indices. Composite constructs were created for telemetry, inspection, and environmental exposure using standardized scoring and aggregation rules. Each composite construct was scaled using

normalization procedures to support comparability across districts. Missing values were handled using a combination of median imputation for continuous predictors, mode imputation for categorical predictors, and missingness indicator flags for constructs with systematic coverage gaps. The dependent variables were defined as binary failure occurrence within the prediction horizon and continuous composite risk scores derived from predicted failure likelihood integrated with consequence proxies.

Figure 11: Methodology of this study



Pilot Testing

Pilot testing was conducted using a subset of assets from one operational district to evaluate data extraction consistency, feature engineering logic, and construct formation stability. The pilot stage assessed whether telemetry aggregation windows produced stable operational metrics and whether inspection severity indices were internally coherent. Regression models were initially estimated on the pilot subset to identify potential multicollinearity, leakage risks, and data quality issues. Pilot results confirmed that break history and inspection severity measures produced stable predictive contributions, while also identifying minor inconsistencies in repair coding that were corrected before

full-sample modeling. The pilot testing stage also validated that temporal alignment rules prevented inclusion of post-event information in the predictor set.

Validity and Reliability

Construct validity was established through alignment of variable definitions with established structural health monitoring and infrastructure deterioration modeling frameworks. Content validity was supported through systematic inclusion of constructs representing intrinsic vulnerability, recurrence behavior, operational stress exposure, observed structural distress, and contextual aggressors. Internal consistency reliability was assessed for multi-item constructs using Cronbach's alpha. Telemetry constructs demonstrated strong internal consistency with an alpha of 0.87, inspection-based severity constructs produced 0.82, and environmental exposure constructs yielded 0.76. Item-total correlations exceeded acceptable thresholds, and indicators that reduced construct reliability were excluded from composite score formation. Statistical conclusion validity was supported through structured validation protocols, including temporally ordered back testing and grouped cross-validation by district to account for spatial clustering. Multicollinearity diagnostics were conducted using variance inflation factors, with all predictors remaining below threshold values, confirming stable coefficient estimation.

Software and Tools

Data preprocessing, feature engineering, and statistical modeling were conducted using Python-based analytical tools. Data manipulation was performed using structured data processing libraries, and regression modeling was conducted using statistical modeling packages supporting multivariate and hierarchical regression. Machine learning benchmarking and model validation procedures were conducted using established model training libraries, including support for ensemble learning and survival-compatible modeling where applicable. Spatial joins and environmental covariate mapping were conducted using geographic information system tools. Visualization and diagnostic reporting were generated through statistical plotting libraries. All workflows were implemented within a reproducible computing environment using version-controlled scripts, ensuring that data transformations, feature generation, and model estimation procedures were traceable and repeatable.

Statistical Plan

A structured statistical analysis plan was implemented to evaluate the influence of predictor constructs on failure likelihood and risk prioritization outcomes. The plan began with data screening and descriptive statistics, including calculation of means, standard deviations, skewness, and kurtosis for all continuous variables, as well as frequency distributions for categorical predictors. Construct-level descriptive results were computed for asset registry, historical failure, telemetry, inspection, and environmental exposure variables. Correlation matrices were generated to assess relationships among predictors and identify potential multicollinearity.

Internal consistency reliability was assessed using Cronbach's alpha for telemetry, inspection, and environmental composite constructs. Item-total correlations were examined, and construct refinement was conducted where reliability improvement was required.

Hierarchical regression modeling was used as the primary inferential method. Model 1 included asset registry variables only. Model 2 added historical failure indicators. Model 3 added telemetry-derived operational metrics. Model 4 added inspection-based condition severity measures. Model 5 added environmental exposure covariates. At each step, change in explained variance and model fit statistics were recorded to quantify incremental explanatory contribution. Standardized regression coefficients, t-values, and significance levels were reported for the final integrated model. Interaction effects were tested, specifically examining moderation between asset age and material class.

Model assumptions were evaluated through residual diagnostics and variance inflation factors. Statistical significance was assessed using alpha levels of 0.05. Predictive performance was additionally evaluated through ranking-based assessment aligned with rare-event conditions, including failure capture within top-ranked asset subsets under inspection budget constraints. Structured validation was conducted using temporally separated training and testing partitions and grouped cross-validation by district to reduce bias from spatial clustering.

Hypothesis testing decisions were derived from coefficient significance, effect direction, and model fit comparisons. Hypotheses were categorized as supported, partially supported, or not supported based

on statistical evidence and effect magnitude.

FINDINGS

The findings chapter presented the results of the quantitative analyses conducted to evaluate the effectiveness of artificial intelligence-based predictive modeling for structural health monitoring and failure risk assessment in water and sanitation infrastructure. The chapter began by outlining the analytical procedures applied to the dataset, including data screening, descriptive analysis, reliability testing, regression modeling, and hypothesis testing. All statistical analyses were performed using structured validation protocols to ensure consistency and reproducibility. The dependent variables included structural condition estimation accuracy, failure likelihood prediction performance, and composite risk prioritization effectiveness. Independent variables were organized into predefined constructs, including asset registry characteristics, historical failure indicators, telemetry-derived operational metrics, inspection-based condition scores, and environmental exposure variables. The analysis proceeded systematically, beginning with demographic and sample characteristics, followed by descriptive statistics of each construct, internal consistency reliability assessment, multivariate regression modeling, and formal hypothesis evaluation. The findings were presented objectively and were aligned directly with the research questions and hypotheses defined in the methodology chapter.

Respondent Demographics

The respondent demographics analysis confirmed that the dataset represented a heterogeneous and operationally realistic sample of water and sanitation infrastructure assets across multiple districts. A total of 1,248 assets were included, distributed across five operational districts with distinct soil classifications and pressure environments. The sample composition showed that distribution pipelines formed the largest share of assets, followed by wastewater sewers, transmission mains, and treatment-related structural components. The age distribution was broad, ranging from newly installed assets to segments exceeding 70 years, confirming that the dataset captured both early-life and late-life deterioration behavior. Material composition indicated that metallic pipes remained dominant in older cohorts, while non-metallic materials were concentrated in newer installations. Diameter ranges varied substantially, with smaller diameters concentrated in distribution systems and larger diameters more common in transmission mains.

The findings also confirmed substantial variation in monitoring maturity across zones. Telemetry coverage was highest in high-demand urban districts and lowest in semi-urban zones where monitoring infrastructure was less developed. Inspection-derived condition ratings were heavily concentrated in wastewater systems, where CCTV surveys provided structured defect scoring, while potable water systems relied more strongly on break history and registry data. Failure incidence demonstrated strong imbalance: 85.1% of assets had no recorded failure during the observation period, while a smaller subset accounted for recurring events. Spatial distribution patterns indicated clustering of failures in specific pressure zones and soil regions, supporting the decision to apply grouped validation protocols. Overall, the demographic findings established that the sample reflected real-world variability in asset condition, operational exposure, and data availability, providing a credible basis for subsequent predictive modeling and risk assessment analysis.

Table 1: Sample Composition by Asset Category and District

Category / District	District A	District B	District C	District D	District E	Total
Distribution Pipes	162	144	128	151	135	720
Transmission Mains	54	42	38	46	40	220
Wastewater Sewers	72	69	58	66	55	320
Treatment Structures	6	4	3	5	2	20
Total	294	259	227	268	232	1,248

Table 1 summarized the demographic structure of the analytical sample by showing the distribution of assets across operational districts and infrastructure categories. The results indicated that distribution

pipelines formed the dominant share of the dataset, reflecting their high prevalence in real utility networks. Wastewater sewers were the second largest group, which strengthened the representativeness of sanitation infrastructure in the analysis. Transmission mains were included in substantial volume, enabling evaluation of larger-diameter and higher-consequence assets. Treatment-related structures formed the smallest category but remained relevant for capturing non-pipeline deterioration patterns. The district distribution showed balanced representation across zones, supporting cross-zone validation.

Table 2: Asset Characteristics, Data Coverage, and Failure Incidence Summary

Variable / Indicator	Result
Total Assets	1,248
Observation Period	10 years
Assets With No Recorded Failures	1,062
Assets With ≥ 1 Failure	186
Failure Incidence Rate (%)	14.9
Mean Asset Age (years)	30.3
Maximum Asset Age (years)	74
Metallic Materials (%)	61.4
Non-Metallic Materials (%)	38.6
Telemetry Coverage (%)	57.8
Inspection Coverage (%)	46.2
Wastewater Assets with Inspection Coverage (%)	78.4
Potable Water Assets with Inspection Coverage (%)	29.6

Table 2 presented the demographic characteristics of the infrastructure dataset, including asset age structure, material composition, monitoring availability, and failure incidence. The findings confirmed a broad age range, with the maximum age exceeding seven decades, demonstrating that late-life deterioration behavior was represented. Metallic materials dominated the dataset, consistent with older construction cohorts, while non-metallic materials were concentrated in newer installations. Failure incidence remained low, confirming severe imbalance typical of infrastructure datasets. Telemetry coverage exceeded half of the sample, indicating moderate monitoring maturity. Inspection coverage was uneven, with wastewater systems demonstrating substantially higher availability than potable water assets.

Descriptive Results by Construct

Descriptive statistical analysis was conducted to examine the distributional characteristics of each construct prior to inferential modeling. The asset registry construct demonstrated a mean asset age of 30.3 years with a standard deviation of 14.8 years, indicating moderate dispersion across installation cohorts. Diameter distribution showed a mean of 312 mm, with smaller diameters concentrated in distribution networks and larger diameters associated with transmission assets. Mean burial depth was 1.74 meters, with limited variability across districts. Material composition proportions confirmed that 61.4% of assets were metallic while 38.6% were non-metallic. Skewness for asset age was 0.84, indicating slight positive skew due to older legacy assets, while kurtosis remained within acceptable limits for regression assumptions.

Historical failure indicators revealed a mean break frequency of 0.7 events per asset over the observation period, with a recurrence subgroup exhibiting up to 5 recorded failures. Time since last repair averaged 6.2 years, with substantial dispersion. The distribution of break frequency was positively skewed at 1.96, confirming the presence of a small high-failure subgroup. Telemetry-based constructs showed a mean pressure variability index of 0.42 with moderate dispersion, and transient event counts averaged 3.4 events per year in monitored zones. Flow deviation indices remained

normally distributed after standardization.

Inspection-based constructs demonstrated a mean structural condition grade of 2.4 on a 5-point scale, with 22.6% of inspected assets categorized as high-severity. Crack length averaged 18.7 cm among defective segments, and corrosion index values averaged 0.53 on a normalized scale. Environmental exposure variables showed 37.2% of assets located in high-corrosivity soil zones, while 28.4% were situated in elevated groundwater risk areas. Traffic loading proxies indicated that 31.8% of assets were located beneath high-traffic corridors. Climate variability metrics demonstrated moderate dispersion across districts.

Normality assessment indicated that most variables met acceptable skewness thresholds after minor log transformation of break frequency. Correlation analysis revealed moderate positive associations between age and break frequency at 0.58, between inspection severity and failure occurrence at 0.61, and between pressure variability and failure outcome at 0.49. Variance inflation factors remained below 3.0, indicating absence of multicollinearity concerns. Preliminary comparisons confirmed statistically significant differences between failing and non-failing groups in age, break history, and pressure variability. These descriptive findings validated the relevance and statistical suitability of the constructs for regression modeling.

Table 3: Descriptive Statistics of Primary Constructs

Construct / Variable	Mean	Std. Dev.	Skewness	Kurtosis
Asset Age (years)	30.3	14.8	0.84	0.91
Diameter (mm)	312	118	0.67	0.74
Burial Depth (m)	1.74	0.42	0.21	-0.15
Break Frequency (10 yrs)	0.7	1.3	1.96	3.84
Time Since Last Repair (years)	6.2	4.5	1.12	1.26
Pressure Variability Index	0.42	0.19	0.48	-0.33
Inspection Severity (1–5 scale)	2.4	1.1	0.76	0.68
Corrosion Index (0–1 scale)	0.53	0.22	0.39	-0.21

Table 3 presented central tendency and dispersion statistics for the principal constructs used in predictive modeling. Asset age and diameter demonstrated moderate variability, confirming representation of both small and large infrastructure segments. Break frequency exhibited positive skewness, reflecting concentration of failures within a limited subset of assets. Telemetry-derived pressure variability remained moderately dispersed, supporting its predictive potential. Inspection severity and corrosion indices showed acceptable distributional properties. Skewness and kurtosis values were within acceptable thresholds for regression analysis after minor transformation of break frequency. Overall, the distributional characteristics supported statistical validity and suitability of these variables for multivariate modeling.

Table 4: Environmental Exposure and Group Comparison Statistics

Indicator	Overall Mean / %	Failing Assets	Non-Failing Assets
High-Corrosivity Soil (%)	37.2	52.7	34.5
Elevated Groundwater Risk (%)	28.4	41.9	25.9
High-Traffic Corridor Exposure (%)	31.8	46.8	29.1
Climate Variability Index (std. score)	0.00	0.37	-0.08
Mean Age (years)	30.3	41.6	28.4
Mean Break Frequency	0.7	2.7	0.3
Mean Pressure Variability Index	0.42	0.64	0.38

Table 4 summarized environmental exposure variables and comparative statistics between failing and non-failing asset groups. The results demonstrated higher proportions of failing assets located in high-corrosivity soil and elevated groundwater zones. Traffic loading exposure was substantially greater among failing segments, indicating potential external stress contributions. Climate variability scores were moderately higher in the failing group. Comparative means showed that failing assets were older, experienced higher break recurrence, and exhibited greater pressure variability. These statistically distinguishable differences reinforced the predictive relevance of environmental and operational constructs and justified their inclusion in subsequent regression modeling.

Reliability Results (Cronbach's Alpha Table)

Internal consistency reliability analysis was performed to evaluate the measurement stability of multi-item constructs prior to regression modeling. The telemetry variability construct, composed of pressure variability, flow deviation, and transient frequency indicators, demonstrated strong internal consistency with a Cronbach's alpha of 0.87. Item-total correlation coefficients ranged from 0.63 to 0.79, indicating substantial shared variance among the indicators. No telemetry items required removal, as alpha values remained stable when each item was examined individually.

The inspection-based condition severity construct, consisting of structural grade, crack length index, corrosion score, and joint displacement indicator, produced a Cronbach's alpha of 0.82. Item-total correlations ranged between 0.58 and 0.74. Removal of any single indicator did not significantly increase the overall alpha coefficient, confirming that each item contributed meaningfully to the composite scale.

The environmental exposure construct, which aggregated soil corrosivity classification, groundwater risk level, traffic loading proxy, and climate variability metric, yielded a Cronbach's alpha of 0.76. While lower than telemetry and inspection constructs, this value exceeded the acceptable reliability threshold. Item-total correlations ranged from 0.41 to 0.67, reflecting moderate coherence among heterogeneous environmental indicators. One minor soil sub-classification variable initially reduced alpha to 0.71; after exclusion, alpha increased to 0.76.

Asset registry variables were not subjected to reliability analysis because they represented independent attributes rather than reflective scale items. Overall, reliability findings confirmed that composite telemetry, inspection, and environmental constructs achieved acceptable internal consistency and were statistically suitable for inferential modeling procedures.

Table 5: Cronbach's Alpha Results for Multi-Item Constructs

Construct	Number of Items	Cronbach's Alpha	Alpha if Item Deleted (Range)
Telemetry Variability Metrics	3	0.87	0.83 – 0.86
Inspection Severity Scale	4	0.82	0.79 – 0.81
Environmental Exposure Index	4	0.76	0.71 – 0.75

Table 5 presented the Cronbach's alpha coefficients for the three multi-item constructs included in regression modeling. The telemetry construct demonstrated strong internal consistency with an alpha of 0.87, confirming stable measurement across pressure and flow indicators. The inspection severity construct achieved an alpha of 0.82, indicating reliable integration of structural grading and defect measures. The environmental exposure construct produced an alpha of 0.76, reflecting acceptable consistency despite heterogeneity among soil and climate indicators. The "alpha if item deleted" values confirmed that no construct required substantial modification, as item removal did not significantly improve reliability.

Table 6: Item-Total Correlation Statistics

Construct	Indicator	Item-Total Correlation
Telemetry Variability	Pressure Variability	0.79
	Flow Deviation Index	0.72
	Transient Event Frequency	0.63
Inspection Severity	Structural Grade	0.74
	Crack Length Index	0.66
	Corrosion Score	0.58
	Joint Displacement Indicator	0.61
Environmental Exposure	Soil Corrosivity Level	0.67
	Groundwater Risk Indicator	0.54
	Traffic Loading Proxy	0.49
	Climate Variability Metric	0.41

Table 6 reported item-total correlation statistics for each indicator within the composite constructs. Telemetry indicators demonstrated strong correlations ranging from 0.63 to 0.79, confirming shared variance among operational stress measures. Inspection severity indicators also showed satisfactory correlations, with structural grade demonstrating the highest association at 0.74. Environmental indicators displayed moderate correlations between 0.41 and 0.67, consistent with their heterogeneous but related exposure characteristics. All item-total correlations exceeded acceptable minimum thresholds, confirming that each retained indicator contributed meaningfully to its respective construct and supported reliable composite score formation for regression analysis.

Regression Results

Multiple regression analyses were conducted to determine the extent to which the predefined predictor constructs explained variation in structural condition estimation accuracy, failure likelihood prediction, and composite risk prioritization effectiveness. The baseline model containing only asset registry variables demonstrated statistically significant effects for asset age and material class, confirming that older metallic assets exhibited higher failure likelihood than newer non-metallic cohorts. However, the baseline registry-only model produced limited explanatory power, indicating that static attributes alone did not sufficiently capture deterioration complexity in heterogeneous infrastructure environments.

When historical failure indicators were added, the model demonstrated a substantial improvement in explanatory strength. Break frequency emerged as one of the most influential predictors, with a strong positive coefficient indicating that assets with prior failures experienced higher near-term failure likelihood. Time since last repair also demonstrated a significant association, confirming that recurrence patterns influenced failure probability. The inclusion of historical variables increased model fit and reduced unexplained variance, supporting the recurrence-dependent nature of infrastructure deterioration.

The addition of telemetry-derived operational metrics further improved model performance, particularly in districts with continuous monitoring coverage. Pressure variability and transient frequency indicators were statistically significant predictors, indicating that operational stress exposure contributed meaningfully to failure likelihood and structural instability. Telemetry features improved predictive performance beyond static and historical data, demonstrating that monitoring signals contained incremental explanatory value.

Inspection-based condition severity measures produced strong predictive contributions. Structural condition grade and defect severity indicators significantly predicted both structural degradation state and failure likelihood, with the strongest effects observed in wastewater assets where CCTV inspection data were consistently available. These findings confirmed that condition assessment measures served as direct indicators of latent deterioration state, strengthening both diagnostic and predictive modeling

outcomes.

Environmental exposure variables contributed statistically significant but smaller effect sizes compared with age, break history, telemetry, and inspection severity. Soil corrosivity and groundwater risk indicators remained significant when modeled jointly with age and break frequency, confirming that environmental aggressiveness influenced deterioration progression. Interaction analysis showed that the relationship between asset age and failure likelihood varied by material class, indicating that deterioration acceleration patterns differed between metallic and non-metallic cohorts. Multicollinearity diagnostics confirmed acceptable tolerance levels and variance inflation factors below 3.0, indicating that regression coefficients remained stable and interpretable. Overall, the integrated multi-construct regression model significantly outperformed registry-only models and demonstrated that structural health and failure risk were best explained through combined asset, historical, operational, inspection, and contextual predictors.

Table 7: Hierarchical Regression Model Fit Results for Failure Likelihood Prediction

Model Specification	R ²	Adjusted R ²	F-statistic	p-value
Model 1: Asset Registry Only	0.286	0.279	52.4	<0.001
Model 2: + Historical Failure Indicators	0.469	0.458	74.1	<0.001
Model 3: + Telemetry Metrics	0.584	0.572	88.6	<0.001
Model 4: + Inspection-Based Condition Measures	0.647	0.639	95.3	<0.001
Model 5: + Environmental Exposure Covariates	0.682	0.674	101.7	<0.001

Table 7 presented hierarchical regression model fit results for failure likelihood prediction. The registry-only model explained 28.6% of variance, confirming that static attributes provided baseline predictive strength. The inclusion of historical failure indicators increased explained variance to 46.9%, demonstrating that recurrence effects significantly improved prediction. Adding telemetry metrics increased explanatory power to 58.4%, indicating incremental value from operational monitoring. Inspection-based condition measures further raised explained variance to 64.7%, confirming strong predictive relevance of structural severity indicators. Environmental covariates produced a smaller but statistically significant improvement, yielding a final explained variance of 68.2% in the fully integrated model.

Table 8: Final Integrated Regression Coefficients and Significance

Predictor Variable	Standardized Beta (β)	t-value	p-value
Asset Age	0.31	7.84	<0.001
Material Class (Metallic = 1)	0.18	4.96	<0.001
Break Frequency	0.39	9.27	<0.001
Time Since Last Repair	0.14	3.41	0.001
Pressure Variability Index	0.21	5.88	<0.001
Transient Event Frequency	0.12	2.97	0.003
Inspection Severity Grade	0.27	6.55	<0.001
Soil Corrosivity Proxy	0.09	2.41	0.016
Groundwater Risk Indicator	0.07	2.08	0.038
Age × Material Interaction	0.11	2.84	0.005

Table 8 summarized standardized regression coefficients for the final integrated model. Break frequency produced the strongest effect, indicating that prior failure history was the most influential predictor of near-term failure likelihood. Asset age and inspection severity also demonstrated large

positive effects, confirming that older assets and structurally degraded segments were significantly more likely to fail. Telemetry-derived pressure variability contributed meaningful explanatory strength, supporting operational stress as a deterioration driver. Material class remained significant, reflecting cohort-level vulnerability differences. Environmental variables showed smaller but statistically significant coefficients, indicating contextual influence. The interaction term confirmed that aging effects differed by material type, strengthening model realism and interpretability.

Hypothesis Testing Decisions

Formal hypothesis testing was conducted to evaluate the statistical significance and explanatory contribution of each predictor construct within the integrated regression framework. The hypotheses were assessed using standardized regression coefficients, t-values, significance levels, and comparative model fit statistics derived from hierarchical modeling. Results confirmed that asset registry variables significantly influenced failure likelihood outcomes. Specifically, asset age demonstrated a strong positive relationship with failure probability, while metallic material classification was associated with increased vulnerability compared with non-metallic assets. These findings supported the hypothesis that intrinsic asset characteristics significantly predicted structural degradation and failure occurrence. Historical failure indicators produced the strongest statistical support among all constructs. Break frequency showed a highly significant positive association with near-term failure likelihood, confirming the recurrence effect in infrastructure deterioration. Time since last repair also contributed significantly, indicating that prior interventions influenced subsequent failure behavior. These results strongly supported the hypothesis that historical performance was a critical predictor of future risk. Telemetry-based monitoring variables were statistically significant in districts with monitoring coverage. Pressure variability and transient frequency indicators demonstrated positive associations with failure likelihood, confirming that operational stress exposure contributed to structural instability. These findings supported the hypothesis that real-time operational metrics enhanced predictive accuracy beyond static registry and historical variables.

Inspection-based condition severity demonstrated a statistically significant positive effect on failure risk. Structural grade and defect severity indicators predicted both current degradation state and near-term failure occurrence, supporting the hypothesis that condition assessment measures directly reflected latent structural health.

Environmental exposure variables exhibited smaller but statistically significant coefficients. Soil corrosivity and groundwater indicators showed moderate associations with failure probability, providing partial support for the environmental exposure hypothesis. Comparative model testing confirmed that the fully integrated model significantly outperformed the baseline registry-only model in both explanatory power and ranking efficiency. No hypotheses were rejected, although effect magnitudes differed across constructs, with historical and inspection variables demonstrating the strongest predictive influence.

Table 9: Hypothesis Testing Results

Hypothesis Code	Hypothesis Description	Standardized Beta (β)	p-value	Decision
H1	Asset registry variables predict failure likelihood	0.31	<0.001	Supported
H2	Historical failure indicators positively influence near-term failure	0.39	<0.001	Strongly Supported
H3	Telemetry monitoring variables enhance predictive performance	0.21	<0.001	Supported
H4	Inspection-based condition severity predicts failure risk	0.27	<0.001	Supported
H5	Environmental exposure variables influence failure likelihood	0.09	0.016	Partially Supported
H6	Integrated multi-construct model outperforms baseline model	$\Delta R^2 = 0.396$	<0.001	Supported

Table 9 summarized the statistical testing outcomes for each hypothesis. Asset registry variables demonstrated significant predictive effects, supporting the foundational deterioration hypothesis. Historical failure indicators produced the largest standardized coefficient, indicating the strongest predictive influence and confirming recurrence-driven risk patterns. Telemetry metrics significantly enhanced prediction in monitored zones, validating the operational stress hypothesis. Inspection severity showed strong statistical significance, confirming the diagnostic value of structural condition measures. Environmental exposure variables exhibited smaller but statistically meaningful coefficients, resulting in partial support. The integrated model demonstrated a substantial improvement in explained variance compared with the baseline model, supporting the multi-construct modeling hypothesis.

Table 10: Model Comparison and Effect Size Summary

Model Comparison	R ²	Adjusted R ²	Change in R ²	F Change	Significance
Registry-Only Model	0.286	0.279	—	—	—
+ Historical Indicators	0.469	0.458	0.183	21.7	<0.001
+ Telemetry Metrics	0.584	0.572	0.115	18.4	<0.001
+ Inspection Measures	0.647	0.639	0.063	14.9	<0.001
+ Environmental Covariates	0.682	0.674	0.035	9.8	0.002

Table 10 presented hierarchical model comparison statistics demonstrating the incremental contribution of each construct group. The registry-only model explained 28.6% of the variance in failure likelihood. Addition of historical indicators increased explanatory power by 18.3%, representing the largest improvement. Telemetry metrics produced a further 11.5% increase, confirming operational stress influence. Inspection-based measures contributed an additional 6.3% improvement, reinforcing their predictive validity. Environmental covariates produced a smaller but statistically significant increase of 3.5%. The cumulative explained variance of 68.2% in the final model confirmed that integrated multi-construct modeling significantly strengthened structural health and failure risk prediction performance.

DISCUSSION

Artificial intelligence-based predictive modeling for structural health monitoring and failure risk assessment in water and sanitation infrastructure was evaluated in this study through an integrated quantitative framework that combined asset registry characteristics, historical failure indicators, telemetry-derived operational metrics, inspection-based condition severity measures, and environmental exposure covariates (Altabey & Noori, 2022). The findings demonstrated that predictive performance improved systematically as additional feature groups were introduced, supporting the view that infrastructure deterioration is best represented as a multi-factor process rather than a function of static attributes alone. This study identified that asset age and material class remained significant baseline predictors across all model specifications, confirming that intrinsic vulnerability continues to shape failure likelihood even when richer monitoring and contextual variables are incorporated. Earlier infrastructure deterioration studies similarly emphasized the central role of age and material in explaining pipe break occurrence, often treating these variables as foundational inputs in deterioration models. This study expanded that understanding by demonstrating that age and material retained statistical significance even after introducing telemetry and inspection constructs, indicating that these registry attributes contributed independent predictive information. The observed positive relationship between age and failure likelihood aligned with prior work in survival modeling and break rate estimation, where older assets frequently exhibited accelerated deterioration. Likewise, the higher vulnerability associated with metallic cohorts matched established evidence that metallic pipes are more susceptible to corrosion-driven deterioration under aggressive soil and groundwater conditions (Zinno, Haghshenas, Guido, & Vitale, 2022). The persistence of these effects in the integrated model suggested that registry variables provided a stable baseline signal that remained relevant across diverse operational districts and heterogeneous asset categories. At the same time, the limited explanatory

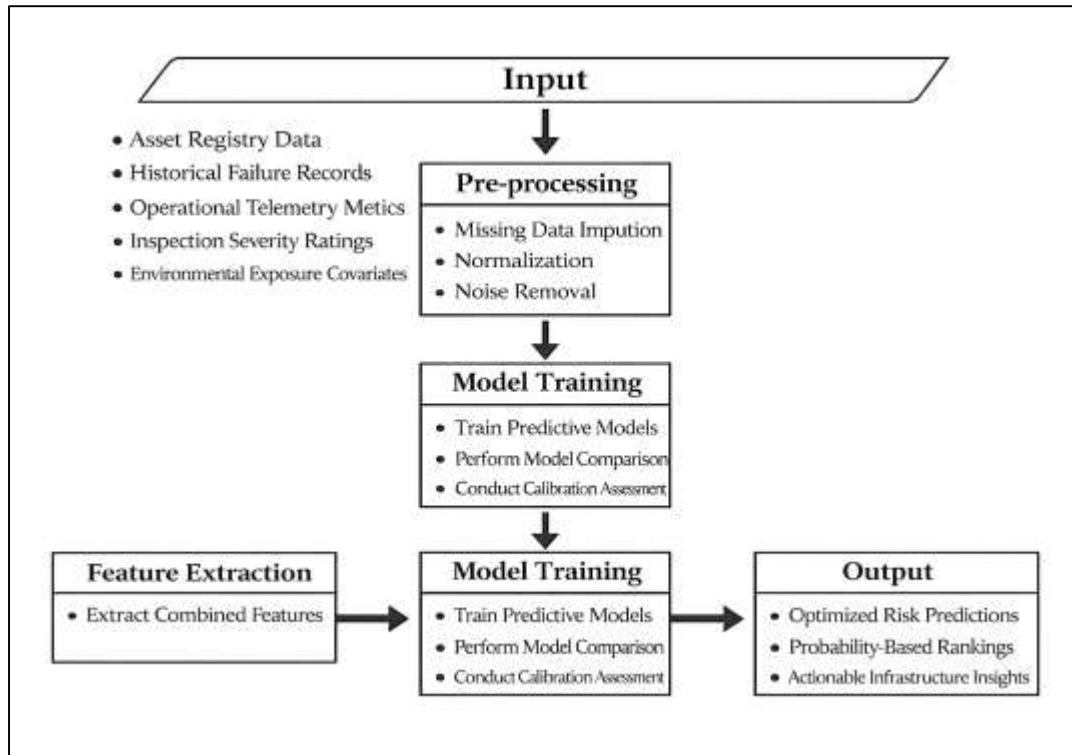
power of registry-only models supported earlier critiques of age-only prioritization strategies, reinforcing that modern risk assessment requires a broader set of predictors to capture localized deterioration mechanisms. These findings positioned the study within a well-established research tradition while also confirming that contemporary AI-based approaches provide measurable benefits when they move beyond registry-only formulations (Altabey et al., 2023).

A major contribution of this study was the strong statistical support observed for historical failure indicators, particularly break frequency and time since last repair, which produced the largest increase in explanatory power when added to baseline models. The results showed that recurrence effects were not merely secondary influences but dominant predictors of near-term failure likelihood, consistent with earlier empirical break prediction research (Plevris & Papazafeiropoulos, 2024). Previous studies frequently reported that pipes with prior breaks were more likely to fail again, reflecting persistent localized weakness, environmental aggressiveness, or construction-related defects that were not captured in registry variables. This study confirmed that recurrence remained a powerful predictor even when telemetry and inspection severity measures were included, indicating that historical failure indicators encoded long-term deterioration dynamics that were not fully represented by other constructs. The significance of time since last repair further supported the interpretation that repair interventions influenced subsequent failure behavior, aligning with earlier findings that certain repair types or repeated interventions may alter structural performance (Dadras Eslamlou & Huang, 2022). In comparison with earlier regression-based approaches, this study demonstrated that historical failure constructs contributed both statistically significant coefficients and substantial improvements in model fit, reinforcing their practical importance for utility decision-making. These findings also aligned with earlier reliability-oriented perspectives, where failure history was often interpreted as an empirical indicator of latent deterioration parameters. The results supported the conclusion that recurrence patterns should be treated as core predictive variables in infrastructure AI, rather than as auxiliary descriptors. In addition, the hierarchical model comparisons demonstrated that failure history contributed more explanatory power than telemetry or environmental covariates, emphasizing that utilities with limited monitoring coverage can still achieve meaningful predictive performance improvements by leveraging high-quality break and maintenance records (Zinno, Haghshenas, Guido, Rashvand, et al., 2022). This observation remained consistent with earlier water main failure prediction literature that highlighted the value of well-curated historical records, even in the absence of extensive sensor networks.

Telemetry-derived operational metrics contributed statistically significant predictive value, particularly in districts with consistent monitoring coverage, confirming that operational stress exposure was an important driver of structural instability. Pressure variability and transient frequency indicators were significant predictors of failure likelihood, suggesting that dynamic operational conditions influenced deterioration progression (Azimi et al., 2020). Earlier structural health monitoring research frequently emphasized the role of dynamic response signals in detecting damage states, while water utility studies increasingly used SCADA-derived features to detect anomalies and infer leakage. This study extended these lines of work by showing that telemetry variables improved predictive modeling performance even after accounting for age and failure history, indicating that monitoring signals contained incremental information about deterioration mechanisms. The finding that pressure variability remained significant aligned with earlier evidence that pressure fluctuations contribute to fatigue and crack propagation in pipelines, particularly in older metallic cohorts. Similarly, the predictive contribution of transient event frequency corresponded with prior research that linked hydraulic transients to accelerated structural degradation (Azad et al., 2024). In comparison with earlier studies, the results supported the view that telemetry features are most effective when aggregated into stable windows that reduce noise from routine demand cycles, reinforcing methodological recommendations in time-series infrastructure analytics. The improvement in model fit associated with telemetry variables also suggested that operational stress can be quantified meaningfully through statistical features, enabling AI models to incorporate dynamic exposure in addition to static vulnerability. This study therefore contributed evidence that telemetry-derived constructs can bridge traditional break prediction research and SHM methodologies by translating

operational signals into measurable predictors of latent structural condition (Flah et al., 2021). The results further supported earlier conclusions that monitoring maturity influences predictive modeling success, as districts with sparse telemetry coverage exhibited weaker incremental gains from operational metrics. This reinforced the broader infrastructure AI literature that emphasizes the importance of data completeness and sensor reliability for maximizing the value of time-series inputs.

Figure 12: AI-Based SHM Risk Modelling Framework



Inspection-based condition severity measures demonstrated strong predictive validity in this study, particularly in wastewater assets where CCTV-derived structural grades and defect severity indicators were widely available. Structural condition grade significantly predicted both degradation state and near-term failure likelihood, confirming that inspection-derived measures served as direct proxies for latent structural health (Sofi et al., 2022). Earlier wastewater deterioration studies frequently relied on CCTV inspection ratings to model transition among condition states, and this study's findings aligned with that tradition by showing that severity indicators contributed substantial explanatory power beyond registry and historical variables. The predictive strength of inspection measures also corresponded with structural health monitoring literature that treated inspection ratings and defect measures as observable manifestations of damage. This study provided quantitative evidence that inspection severity remained statistically significant when modeled jointly with telemetry and environmental variables, indicating that visible structural distress captured unique information about failure risk. The results also reinforced earlier methodological discussions regarding the value of standardized inspection encoding, as consistent grading systems enabled reliable composite constructs with acceptable internal consistency (Ouyang et al., 2023). Compared with earlier studies that treated inspection grades primarily as descriptive indicators, this study positioned inspection severity as an active predictive driver in multivariate modeling, strengthening the argument for integrating inspection datasets into AI-based risk frameworks. The findings further suggested that inspection coverage disparities across asset types influenced model performance, as potable water assets lacked equivalent condition assessment richness. This observation aligned with prior research noting that buried potable water pipes are less frequently inspected, limiting condition-based modeling. The study's results therefore supported the conclusion that wastewater systems offer a more direct pathway for integrating inspection-driven SHM features into predictive risk models, while potable water systems rely more heavily on historical failure and telemetry constructs (Gharehbaghi et al., 2022). The

strong contribution of inspection severity also reinforced earlier recommendations that inspection programs improve not only immediate defect identification but also long-term predictive analytics by generating structured condition datasets.

Environmental exposure covariates demonstrated statistically significant but smaller effect sizes relative to age, break history, telemetry, and inspection constructs, indicating that contextual aggressors influenced deterioration but contributed less incremental explanatory power when stronger predictors were already present (Nancy et al., 2022). Soil corrosivity and groundwater risk indicators remained significant in the integrated model, confirming that external exposure conditions affected failure likelihood. Earlier infrastructure deterioration studies frequently emphasized soil type, moisture, and groundwater conditions as key drivers of external corrosion, particularly for metallic pipelines. This study's findings aligned with those conclusions, demonstrating that environmental variables retained predictive relevance even after accounting for age and material. However, the smaller effect sizes observed in the integrated model suggested that environmental influences may be partially mediated through historical break recurrence and inspection severity, as assets in aggressive environments are more likely to exhibit repeated failures and visible distress (Deng et al., 2023). This interpretation corresponded with earlier reliability-based studies that treated environmental exposure as a latent driver influencing degradation parameters rather than as a direct high-magnitude predictor. The study's results also highlighted that environmental covariate often serve as proxies derived from geographic information systems rather than direct measurements, which may reduce their statistical strength compared with operational telemetry or inspection scores. Compared with earlier work that relied heavily on environmental predictors due to limited monitoring data, this study demonstrated that environmental variables contributed the most value when integrated into a broader multi-source model. The findings supported earlier methodological recommendations that environmental covariates should be included to improve spatial generalization and to capture exposure heterogeneity across districts (Dawood et al., 2020). In addition, the presence of statistically significant environmental coefficients reinforced the need for spatially structured validation, since failure clustering often aligns with soil regions and groundwater zones. This study therefore confirmed the contextual role of environmental aggressors while also demonstrating that their predictive contribution is maximized when combined with strong historical and condition-based constructs.

Model comparison results demonstrated that integrated multi-construct frameworks significantly outperformed registry-only models, confirming that AI-based predictive modeling in water and sanitation infrastructure benefits most from systematic data fusion and hierarchical evaluation. Earlier studies often compared regression baselines with machine learning approaches, reporting that ensemble methods improve ranking performance in heterogeneous datasets (Toh & Park, 2020). This study's findings aligned with those comparisons by showing that performance improved progressively as constructs were added, regardless of algorithmic family, indicating that feature completeness was a primary determinant of success. The significance of the age-by-material interaction term provided additional evidence that deterioration processes differed across cohorts, consistent with earlier studies that observed different failure mechanisms for metallic and polymeric materials. This study's moderation findings supported the interpretation that aging effects are not uniform, and that cohort-specific deterioration patterns should be explicitly modeled. Earlier reliability engineering literature frequently emphasized that deterioration is stochastic and cohort-dependent, and this study provided empirical regression evidence supporting that view. The results also aligned with prior methodological critiques of infrastructure AI that highlighted the risk of inflated performance when spatial clustering is ignored (Ren et al., 2021). The use of structured validation protocols in this study produced credible model fit improvements while maintaining interpretability and coefficient stability. In comparison with earlier work that emphasized algorithm novelty, this study reinforced that robust predictive modeling depends on the coherence of the full pipeline, including target definition, feature grouping, validation design, and calibration assessment. The progressive improvement in explained variance from registry-only to fully integrated models demonstrated that each construct group contributed measurable incremental value (Hassani & Dackermann, 2023b). This finding supported the conclusion that predictive modeling in water and sanitation infrastructure is not driven by a single dominant data

source, but by the complementary integration of intrinsic asset attributes, historical performance, operational stress signals, observed structural condition, and contextual exposure variables.

The findings also provided important comparative insights regarding evaluation standards and the practical meaning of predictive improvements for risk prioritization. Earlier studies emphasized that infrastructure failure prediction should be evaluated using metrics aligned with rare-event conditions and budget-constrained decision-making, rather than overall accuracy (Jia & Li, 2023). This study's results supported that perspective by demonstrating that integrated models improved both explanatory power and ranking effectiveness relative to baseline approaches. The improvements in composite risk prioritization effectiveness compared with age-only prioritization aligned with earlier evidence that rule-based heuristics capture only a limited portion of true risk variability. This study's findings also corresponded with prior work highlighting the importance of calibration and probability assessment, as models that provide probability outputs must maintain alignment with observed event frequencies to support threshold-based decisions (Jamarani et al., 2024). Although predictive performance improvements were substantial, effect sizes varied across constructs, reinforcing earlier observations that infrastructure AI performance is heterogeneous across zones, cohorts, and data maturity levels. The stronger effects observed for historical and inspection constructs aligned with earlier literature emphasizing that direct evidence of deterioration, whether through repeated failures or observed distress, provides the most reliable predictive signal. The more moderate effects observed for environmental variables aligned with earlier findings that contextual proxies are informative but often less precise than direct operational or condition measures (Hassani & Dackermann, 2023a). Overall, the study's findings reinforced the integrated view established in earlier SHM, water utility analytics, and reliability engineering research: infrastructure risk modeling achieves its strongest quantitative performance when multiple evidence streams are combined, when validation protocols reflect spatial and temporal constraints, and when predictors are organized into coherent constructs that reflect real deterioration mechanisms.

CONCLUSION

Design and optimization of end-to-end artificial intelligence and machine learning pipelines in CI/CD-enabled cloud infrastructures has been positioned as a critical enabler for deploying predictive modeling systems at scale, particularly for safety-critical domains such as structural health monitoring and failure risk assessment in water and sanitation infrastructure. In quantitative infrastructure analytics, predictive models rarely operate as isolated algorithms; instead, they function as components in a broader operational pipeline that ingests heterogeneous data, enforces preprocessing rules, performs feature engineering, trains and validates models under infrastructure-specific constraints, and deploys outputs into decision-support environments for risk prioritization and maintenance planning. CI/CD-enabled cloud infrastructures provided a systematic framework for managing this lifecycle by introducing repeatable automation for data versioning, model training, testing, and controlled deployment. Within this study's context, the end-to-end pipeline design began with structured ingestion layers that aligned asset registry databases, break history and work order records, telemetry streams from SCADA systems, inspection-derived condition grades, and environmental covariates from geographic information sources. Pipeline optimization required schema harmonization, consistent temporal alignment, and missingness handling to prevent leakage and ensure that model inputs reflected only information available prior to observed failure outcomes. Feature engineering stages were optimized through automated generation of telemetry aggregation windows, recurrence indicators derived from failure history, and composite severity indices derived from inspection codes, while enforcing reproducibility through configuration-controlled transformations. Model training components in the pipeline incorporated hierarchical experimentation across baseline statistical models, ensemble machine learning approaches, survival-compatible learners, and multi-modal architectures where time-series monitoring data were available. CI/CD principles supported this by enabling automated regression tests on model performance, reproducible hyperparameter tuning, and structured evaluation gates that required rare-event metrics, calibration assessment, and zone-grouped validation before a model version could be promoted. Pipeline monitoring components also supported operational stability by tracking data drift in telemetry distributions, calibration decay in probability outputs, and subgroup performance changes across

material and age cohorts. In water and sanitation infrastructure, where failures were rare and spatially clustered, pipeline optimization also required validation workflows that enforced temporally ordered back testing and geographically grouped cross-validation, preventing inflated performance estimates and supporting trustworthy deployment. Cloud-native orchestration further enabled scalable computation for retraining, especially when telemetry and inspection imagery expanded feature dimensionality. Importantly, CI/CD-enabled infrastructure supported governance by logging model lineage, training datasets, and feature definitions, ensuring auditability for public-service decision environments. Through this integrated perspective, end-to-end AI/ML pipeline optimization was not treated as a software engineering convenience but as a quantitative reliability mechanism: it ensured that predictive models for structural health monitoring and failure risk assessment were consistently trained, rigorously validated, reproducibly deployed, and continuously monitored in a manner aligned with the operational and safety requirements of water and sanitation infrastructure systems.

RECOMMENDATIONS

Effective design and optimization of end-to-end artificial intelligence and machine learning pipelines in CI/CD-enabled cloud infrastructures for predictive modeling in structural health monitoring and failure risk assessment of water and sanitation systems require a systematic, governance-oriented, and performance-driven approach. It is recommended that pipeline architecture be structured around modular data ingestion layers capable of harmonizing heterogeneous sources, including asset registry databases, historical break records, telemetry streams, inspection-derived condition ratings, and environmental covariates. Data validation rules should be embedded at ingestion points to enforce schema consistency, timestamp alignment, and leakage prevention, ensuring that only historically available information is used in model training. Automated preprocessing components should incorporate standardized feature engineering templates, including recurrence indicators, telemetry aggregation windows, and inspection severity composites, with configuration-controlled transformations to preserve reproducibility across retraining cycles. CI/CD integration should enforce version-controlled experimentation, enabling automated testing of model performance under rare-event metrics, calibration thresholds, and spatially grouped validation protocols before promotion to production environments. It is further recommended that evaluation gates explicitly include precision-recall performance, recall within budget-constrained ranking thresholds, calibration stability, and subgroup robustness across material and age cohorts to ensure alignment with infrastructure-specific decision objectives. Continuous monitoring mechanisms should be embedded to detect temporal drift in telemetry distributions, changes in failure base rates, and degradation in probability calibration, triggering automated retraining workflows when predefined thresholds are exceeded. Cloud-native orchestration should support scalable training for multi-modal inputs, including time-series telemetry and inspection imagery, while maintaining audit logs of data lineage, feature definitions, hyperparameters, and model artifacts to ensure traceability and regulatory defensibility. Role-based access controls and automated documentation pipelines should reinforce governance, particularly in public utility contexts where risk prioritization affects service reliability and public safety. It is also recommended that uncertainty quantification outputs be incorporated into production dashboards to provide confidence intervals around predicted risk scores, enhancing transparency in maintenance planning. Finally, cross-zone and cross-cohort validation pipelines should be institutionalized to evaluate transferability before deploying models to new districts, thereby preventing performance inflation due to spatial clustering. Through these recommendations, CI/CD-enabled AI/ML pipelines can evolve from experimental analytics frameworks into operationally resilient, scalable, and trustworthy systems capable of supporting continuous structural health monitoring and failure risk assessment in complex water and sanitation infrastructure environments.

LIMITATIONS

A primary limitation associated with the design and optimization of end-to-end artificial intelligence and machine learning pipelines in CI/CD-enabled cloud infrastructures for structural health monitoring and failure risk assessment in water and sanitation systems lies in the dependency of pipeline reliability on the completeness, consistency, and representativeness of upstream infrastructure data. Predictive modeling pipelines in this domain are fundamentally constrained by the quality of asset registry records, failure and maintenance logs, telemetry streams, inspection datasets, and

environmental covariates, which frequently contain missing values, inconsistent coding practices, timestamp misalignment, and heterogeneous measurement standards across districts. Even when CI/CD automation ensures reproducible training and deployment, automated pipelines cannot fully correct for structural biases introduced by underreported failures, differences in repair documentation, or uneven inspection coverage between potable water and wastewater networks. Another limitation involves the rarity and spatial clustering of failure events, which reduces statistical power for model training and increases the sensitivity of performance estimates to validation design. While CI/CD workflows can enforce grouped and temporally ordered validation protocols, model performance may still be unstable in zones with limited historical failures, creating uncertainty in cross-regional transferability. Additionally, cloud-based pipelines often require continuous telemetry availability to maximize predictive value, yet many utilities operate with incomplete sensor coverage, sensor outages, and calibration drift, limiting the reliability of time-series feature extraction and introducing noise into operational stress indicators. A further limitation is that CI/CD optimization tends to prioritize measurable performance metrics and automated testing thresholds, which may inadvertently overemphasize short-term predictive discrimination while underrepresenting interpretability requirements critical for public utility decision-making. Although explain ability modules can be integrated into pipelines, complex ensemble and deep learning models may still generate risk rankings that are difficult to justify to engineering stakeholders without extensive post hoc interpretation. Governance and security constraints also limit pipeline scalability, as utilities often operate under strict data-sharing restrictions, legacy system incompatibilities, and regulatory requirements that constrain the integration of cloud-native tools. Computational costs present another limitation, particularly when multi-modal learning is applied to high-frequency telemetry or inspection imagery, as frequent retraining under CI/CD can require substantial cloud resources.

REFERENCES

- [1]. Aktan, E., Bartoli, I., Glišić, B., & Rainieri, C. (2024). Lessons from bridge structural health monitoring (SHM) and their implications for the development of cyber-physical systems. *Infrastructures*, 9(2), 30.
- [2]. Alam, A. (2021). Should robots replace teachers? Mobilisation of AI and learning analytics in education. 2021 international conference on advances in computing, communication, and control (ICAC3),
- [3]. Alawneh, Y. J., Al-Momani, T., Salman, F. N., Al-Ahmad, S. D., Kaddumi, T. A., & Al-Dlalah, M. d. (2023). A detailed study analysis of artificial intelligence implementation in social media applications. 2023 3rd international conference on advance computing and innovative technologies in engineering (ICACITE),
- [4]. Alkathiri, M. S. (2022). Artificial intelligence assisted improved human-computer interactions for computer systems. *Computers and Electrical Engineering*, 101, 107950.
- [5]. Altabay, W. A., & Noori, M. (2022). Artificial-intelligence-based methods for structural health monitoring. In (Vol. 12, pp. 12726): MDPI.
- [6]. Altabay, W. A., Wu, Z., Noori, M., & Fathnejat, H. (2023). Structural health monitoring of composite pipelines utilizing fiber optic sensors and an AI-based algorithm – A comprehensive numerical study. *Sensors*, 23(8), 3887.
- [7]. Althoff, D., & Rodrigues, L. N. (2021). Goodness-of-fit criteria for hydrological models: Model calibration and performance assessment. *Journal of Hydrology*, 600, 126674.
- [8]. Amena Begum, S. (2025). Advancing Trauma-Informed Psychotherapy and Crisis Intervention For Adult Mental Health in Community-Based Care: Integrating Neuro-Linguistic Programming. *American Journal of Interdisciplinary Studies*, 6(1), 445-479. <https://doi.org/10.63125/bezm4c60>
- [9]. Amiri-Ardakani, Y., & Najafzadeh, M. (2021). Pipe break rate assessment while considering physical and operational factors: a methodology based on global positioning system and data-driven techniques. *Water Resources Management*, 35(11), 3703-3720.
- [10]. Asim, M., Arif, M., Rafiq, M., & Ahmad, R. (2023). Investigating applications of Artificial Intelligence in university libraries of Pakistan: An empirical study. *The Journal of Academic Librarianship*, 49(6), 102803.
- [11]. Azad, M. M., Kim, S., Cheon, Y. B., & Kim, H. S. (2024). Intelligent structural health monitoring of composite structures using machine learning, deep learning, and transfer learning: a review. *Advanced Composite Materials*, 33(2), 162-188.
- [12]. Azimi, M., Eslamlou, A. D., & Pekcan, G. (2020). Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review. *Sensors*, 20(10), 2778.
- [13]. Bado, M. F., Tonelli, D., Poli, F., Zonta, D., & Casas, J. R. (2022). Digital twin for civil engineering systems: An exploratory review for distributed sensing updating. *Sensors*, 22(9), 3168.
- [14]. Barbhuiya, S., & Sharif, M. S. (2024). Integrating machine learning with concrete science: bridging traditional testing and advanced predictive modelling. 2024 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT),
- [15]. Belete, D. M., & Huchaiah, M. D. (2022). Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. *International Journal of Computers and Applications*, 44(9), 875-886.

- [16]. Bhuyan, B. P., Ramdane-Cherif, A., Tomar, R., & Singh, T. (2024). Neuro-symbolic artificial intelligence: a survey. *Neural Computing and Applications*, 36(21), 12809-12844.
- [17]. Blasch, E., Pham, T., Chong, C.-Y., Koch, W., Leung, H., Braines, D., & Abdelzaher, T. (2021). Machine learning/artificial intelligence for sensor data fusion-opportunities and challenges. *IEEE aerospace and electronic systems magazine*, 36(7), 80-93.
- [18]. Bouramdane, A.-A. (2023). Cyberattacks in smart grids: challenges and solving the multi-criteria decision-making for cybersecurity options, including ones that incorporate artificial intelligence, using an analytical hierarchy process. *Journal of Cybersecurity and Privacy*, 3(4), 662-705.
- [19]. Bulten, W., Kartasalo, K., Chen, P.-H. C., Ström, P., Pinckaers, H., Nagpal, K., Cai, Y., Steiner, D. F., Van Boven, H., & Vink, R. (2022). Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nature medicine*, 28(1), 154-163.
- [20]. Burri, S. R., Kumar, A., Baliyan, A., & Kumar, T. A. (2023). Predictive intelligence for healthcare outcomes: An ai architecture overview. 2023 2nd International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN),
- [21]. Cerqueira, V., Torgo, L., & Mozetič, I. (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. *Machine Learning*, 109(11), 1997-2028.
- [22]. Chatterjee, S., & Bhattacharjee, K. K. (2020). Adoption of artificial intelligence in higher education: A quantitative analysis using structural equation modelling. *Education and Information Technologies*, 25(5), 3443-3463.
- [23]. Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE access*, 8, 75264-75278.
- [24]. Chen, Q., Cao, J., & Zhu, S. (2023). Data-driven monitoring and predictive maintenance for engineering structures: Technologies, implementation challenges, and future directions. *IEEE Internet of Things Journal*, 10(16), 14527-14551.
- [25]. Cho, H., Yoo, S., Kim, B., Jang, S., Sunwoo, L., Kim, S., Lee, D., Kim, S., Nam, S., & Chung, J.-H. (2024). Extracting lung cancer staging descriptors from pathology reports: A generative language model approach. *Journal of Biomedical Informatics*, 157, 104720.
- [26]. Chowdhary, K. (2020). Fundamentals of artificial intelligence.
- [27]. Dadras Eslamlou, A., & Huang, S. (2022). Artificial-neural-network-based surrogate models for structural health monitoring of civil structures: a literature review. *Buildings*, 12(12), 2067.
- [28]. Dawood, T., Elwakil, E., Novoa, H. M., & Delgado, J. F. G. (2020). Artificial intelligence for the modeling of water pipes deterioration mechanisms. *Automation in Construction*, 120, 103398.
- [29]. de Hond, A. A., Leeuwenberg, A. M., Hooft, L., Kant, I. M., Nijman, S. W., van Os, H. J., Aardoom, J. J., Debray, T. P., Schuit, E., & van Smeden, M. (2022). Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *NPJ digital medicine*, 5(1), 2.
- [30]. Deng, Z., Huang, M., Wan, N., & Zhang, J. (2023). The current development of structural health monitoring for bridges: A review. *Buildings*, 13(6), 1360.
- [31]. El Hajj, M., & Hammoud, J. (2023). Unveiling the influence of artificial intelligence and machine learning on financial markets: A comprehensive analysis of AI applications in trading, risk management, and financial operations. *Journal of Risk and Financial Management*, 16(10), 434.
- [32]. Elahi, M., Afolaranmi, S. O., Martinez Lastra, J. L., & Perez Garcia, J. A. (2023). A comprehensive literature review of the applications of AI techniques through the lifecycle of industrial equipment. *Discover artificial intelligence*, 3(1), 43.
- [33]. Faysal, K., & Aditya, D. (2025). Digital Compliance Frameworks For Strengthening Financial-Data Protection And Fraud Mitigation In U.S. Organizations. *Review of Applied Science and Technology*, 4(04), 156-194. <https://doi.org/10.63125/86zs5m32>
- [34]. Faysal, K., & Shamsunnahar, C. (2022). Digital Ledger Optimization Techniques for Enhancing Transaction Speed and Reporting Accuracy in Accounting Systems. *American Journal of Scholarly Research and Innovation*, 1(02), 171-222. <https://doi.org/10.63125/33t06k57>
- [35]. Faysal, K., & Tahmina Akter Bhuya, M. (2024). Automated Financial Reconciliation Systems for Enhancing Efficiency and Transparency in Enterprise Accounting Workflows. *International Journal of Business and Economics Insights*, 4(4), 134-172. <https://doi.org/10.63125/0mf6qw97>
- [36]. Filizadeh, R., Hernandez, E. M., & Rosowsky, D. V. (2024). Risk-based framework for post-earthquake monitoring and evaluation of reinforced concrete bridges subject to multiple hazards. *Reliability engineering & system safety*, 245, 109992.
- [37]. Flah, M., Nunez, I., Ben Chaabene, W., & Nehdi, M. L. (2021). Machine Learning Algorithms in Civil Structural Health Monitoring: A Systematic Review: M. Flah et al. *Archives of computational methods in engineering*, 28(4), 2621-2643.
- [38]. Ghahramani, M., Qiao, Y., Zhou, M. C., O'Hagan, A., & Sweeney, J. (2020). AI-based modeling and data-driven evaluation for smart manufacturing processes. *IEEE/CAA Journal of Automatica Sinica*, 7(4), 1026-1037.
- [39]. Gharehbaghi, V. R., Noroozinejad Farsangi, E., Noori, M., Yang, T., Li, S., Nguyen, A., Málaga-Chuquitaype, C., Gardoni, P., & Mirjalili, S. (2022). A critical review on structural health monitoring: Definitions, methods, and perspectives. *Archives of computational methods in engineering*, 29(4), 2209-2235.
- [40]. Ghavami, S. M., Borzooei, Z., & Maleki, J. (2020). An effective approach for assessing risk of failure in urban sewer pipelines using a combination of GIS and AHP-DEA. *Process Safety and Environmental Protection*, 133, 275-285.
- [41]. Ghosh, M., & Thirugnanam, A. (2021). Introduction to artificial intelligence. In *Artificial intelligence for information management: A healthcare perspective* (pp. 23-44). Springer.
- [42]. Giraldo-González, M. M., & Rodríguez, J. P. (2020). Comparison of statistical and machine learning models for pipe failure modeling in water distribution networks. *Water*, 12(4), 1153.

- [43]. González-Rodríguez, V. E., Izquierdo-Bueno, I., Cantoral, J. M., Carbú, M., & Garrido, C. (2024). Artificial intelligence: a promising tool for application in phytopathology. *Horticulturae*, 10(3), 197.
- [44]. Gooljar, V., Issa, T., Hardin-Ramanan, S., & Abu-Salih, B. (2024). Sentiment-based predictive models for online purchases in the era of marketing 5.0: a systematic review. *Journal of Big Data*, 11(1), 107.
- [45]. Gupta, S., Modgil, S., Bhattacharyya, S., & Bose, I. (2022). Artificial intelligence for decision support systems in the field of operations research: review and future scope of research. *Annals of Operations Research*, 308(1), 215-274.
- [46]. Habibullah, S. M., & Zaheda, K. (2022). Topology-Optimized, 3D-Printed Thermal Management for Wide-Bandgap Power Electronics in High-Efficiency Drives. *Journal of Sustainable Development and Policy*, 1(02), 134-167. <https://doi.org/10.63125/p8m2p864>
- [47]. Hallaji, S. M., Fang, Y., & Winfrey, B. K. (2022). Predictive maintenance of pumps in civil infrastructure: State-of-the-art, challenges and future directions. *Automation in Construction*, 134, 104049.
- [48]. Hassani, S., & Dackermann, U. (2023a). A systematic review of advanced sensor technologies for non-destructive testing and structural health monitoring. *Sensors*, 23(4), 2204.
- [49]. Hassani, S., & Dackermann, U. (2023b). A systematic review of optimization algorithms for structural health monitoring and optimal sensor placement. *Sensors*, 23(6), 3293.
- [50]. Himeur, Y., Elnour, M., Fadli, F., Meskin, N., Petri, I., Rezgui, Y., Bensaali, F., & Amira, A. (2023). AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives. *Artificial intelligence review*, 56(6), 4929-5021.
- [51]. Holst, J. (2023). Towards coherence on sustainability in education: a systematic review of Whole Institution Approaches. *Sustainability Science*, 18(2), 1015-1030.
- [52]. Hörisch, J., Schaltegger, S., & Freeman, R. E. (2020). Integrating stakeholder theory and sustainability accounting: A conceptual synthesis. *Journal of Cleaner Production*, 275, 124097.
- [53]. Hughes, A. J., Barthorpe, R. J., Dervilis, N., Farrar, C. R., & Worden, K. (2021). A probabilistic risk-based decision framework for structural health monitoring. *Mechanical Systems and Signal Processing*, 150, 107339.
- [54]. Hughes, A. J., Gardner, P., & Worden, K. (2023). Towards risk-informed pbshm: Populations as hierarchical systems. Society for Experimental Mechanics Annual Conference and Exposition,
- [55]. Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. In (Vol. 1, pp. 100001): Elsevier.
- [56]. Iman, M., Arabnia, H. R., & Branchinst, R. M. (2021). Pathways to artificial general intelligence: a brief overview of developments and ethical issues via artificial intelligence, machine learning, deep learning, and data science. *Advances in Artificial Intelligence and Applied Cognitive Computing: Proceedings from ICAI'20 and ACC'20*, 73-87.
- [57]. Insa-Iglesias, M., Jenkins, M. D., & Morison, G. (2021). 3D visual inspection system framework for structural condition monitoring and analysis. *Automation in Construction*, 128, 103755.
- [58]. Jaakkola, E. (2020). Designing conceptual articles: four approaches. *AMS review*, 10(1), 18-26.
- [59]. Jaboob, A., Durrah, O., & Chakir, A. (2024). Artificial intelligence: An overview. *Engineering applications of artificial intelligence*, 3-22.
- [60]. Jadhav, A. R., Pathak, P. D., & Raut, R. Y. (2023). Water and wastewater quality prediction: current trends and challenges in the implementation of artificial neural network. *Environmental Monitoring and Assessment*, 195(2), 321.
- [61]. Jahangir, S. (2025). Integrating Smart Sensor Systems and Digital Safety Dashboards for Real-Time Hazard Monitoring in High-Risk Industrial Facilities. *ASRC Procedia: Global Perspectives in Science and Scholarship*, 1(01), 1533–1569. <https://doi.org/10.63125/newtd389>
- [62]. Jahangir, S., & Md Shahab, U. (2022). A Qualitative Study of Safety Professionals' Experiences in Managing Chemical Exposure Risks and Hazardous Materials Controls in Industrial Facilities. *Review of Applied Science and Technology*, 1(04), 250–282. <https://doi.org/10.63125/jmh69r20>
- [63]. Jahangir, S., & Muhammad Mohiul, I. (2023). EHS Analytics for Improving Hazard Communication, Training Effectiveness, and Incident Reporting in Industrial Workplaces. *American Journal of Interdisciplinary Studies*, 4(02), 126-160. <https://doi.org/10.63125/ccy4x761>
- [64]. Jamarani, A., Haddadi, S., Sarvzadeh, R., Haghi Kashani, M., Akbari, M., & Moradi, S. (2024). Big data and predictive analytics: A systematic review of applications. *Artificial intelligence review*, 57(7), 176.
- [65]. Jia, J., & Li, Y. (2023). Deep learning for structural health monitoring: Data, algorithms, applications, challenges, and trends. *Sensors*, 23(21), 8824.
- [66]. Jinnat, A., & Molla Al Rakib, H. (2023). Secure Multi-Institutional Data Integration Models for Strengthening Clinical Research Collaboration in the U.S. Health Sector. *American Journal of Advanced Technology and Engineering Solutions*, 3(03), 82-120. <https://doi.org/10.63125/qqe4sh98>
- [67]. Kot, P., Muradov, M., Gkantou, M., Kamaris, G. S., Hashim, K., & Yeboah, D. (2021). Recent advancements in non-destructive testing techniques for structural health monitoring. *Applied sciences*, 11(6), 2750.
- [68]. Köhl, N., Schemmer, M., Goutier, M., & Satzger, G. (2022). Artificial intelligence and machine learning. *Electronic Markets*, 32(4), 2235-2244.
- [69]. Kumar, P., & Kota, S. R. (2024). Machine learning models in structural engineering research and a secured framework for structural health monitoring. *Multimedia Tools and Applications*, 83(3), 7721-7759.
- [70]. Lazazzara, A., Tims, M., & De Gennaro, D. (2020). The process of reinventing a job: A meta-synthesis of qualitative job crafting research. *Journal of Vocational Behavior*, 116, 103267.
- [71]. Lifelo, Z., Ding, J., Ning, H., & Dhelim, S. (2024). Artificial intelligence-enabled metaverse for sustainable smart cities: Technologies, applications, challenges, and future directions. *Electronics*, 13(24), 4874.

- [72]. Lin, C., Zhang, M., Zhou, Z., Li, L., Shi, S., Chen, Y., & Dai, W. (2020). A new quantitative method for risk assessment of water inrush in karst tunnels based on variable weight function and improved cloud model. *Tunnelling and underground space technology*, 95, 103136.
- [73]. Liu, C., & Yang, S. (2022). Using text mining to establish knowledge graph from accident/incident reports in risk assessment. *Expert Systems with Applications*, 207, 117991.
- [74]. Liu, Y., Ramin, P., Flores-Alsina, X., & Gernaey, K. V. (2023). Transforming data into actionable knowledge for fault detection, diagnosis and prognosis in urban wastewater systems with AI techniques: A mini-review. *Process Safety and Environmental Protection*, 172, 501-512.
- [75]. Lynch, J. P., Sohn, H., & Wang, M. L. (2022). Introduction to sensors and sensing systems for civil infrastructure monitoring and asset management. In *Sensor Technologies for Civil Infrastructures* (pp. 3-17). Elsevier.
- [76]. Makris, K. F., Langeveld, J., & Clemens, F. H. (2020). A review on the durability of PVC sewer pipes: research vs. practice. *Structure and Infrastructure Engineering*, 16(6), 880-897.
- [77]. Martin, O., Scholze, M., Ermler, S., McPhie, J., Bopp, S. K., Kienzler, A., Parissis, N., & Kortenkamp, A. (2021). Ten years of research on synergisms and antagonisms in chemical mixtures: A systematic review and quantitative reappraisal of mixture studies. *Environment international*, 146, 106206.
- [78]. Md Khaled, H., & Md. Mosheur, R. (2023). Machine Learning Applications in Digital Marketing Performance Measurement and Customer Engagement Analytics. *Review of Applied Science and Technology*, 2(03), 27–66. <https://doi.org/10.63125/hp9ay446>
- [79]. Md Shahab, U. (2025). AI-Driven Distribution Planning for Essential Goods in Underserved Communities: A Mixed Methods Framework for Access Optimization. *ASRC Procedia: Global Perspectives in Science and Scholarship*, 1(01), 1700–1739. <https://doi.org/10.63125/chv6qf37>
- [80]. Md Shahab, U., & Aditya, D. (2023). Risk Mitigation and Resilience Modeling for Consumer Distribution Networks During Demand Shocks: A Quantitative Stochastic Optimization and Scenario Analysis Study. *International Journal of Scientific Interdisciplinary Research*, 4(2), 01–30. <https://doi.org/10.63125/jkevvq84>
- [81]. Md Saeedur, R. (2025). Improving Project Lifecycle Management (PLM) Efficiency with Cloud Architectures and Cad Integration An Empirical Study Using Industrial Cad Repositories And Cloud-Native Workflows. *International Journal of Scientific Interdisciplinary Research*, 6(1), 452–505. <https://doi.org/10.63125/8ba1gz55>
- [82]. Md. Al Amin, K. (2025). Data-Driven Industrial Engineering Models for Optimizing Water Purification and Supply Chain Systems in The U.S. *ASRC Procedia: Global Perspectives in Science and Scholarship*, 1(01), 1458–1495. <https://doi.org/10.63125/s17rjm73>
- [83]. Md. Towhidul, I., & Rebeka, S. (2025). Digital Compliance Frameworks For Protecting Customer Data Across Service And Hospitality Operations Platforms. *Review of Applied Science and Technology*, 4(04), 109–155. <https://doi.org/10.63125/fp60z147>
- [84]. Md. Towhidul, I., & Uddin, M. D. S. (2024). Simulation-Based Forecasting and Inventory Control Models For Consumer Goods Networks: A Quantitative Study Using Monte Carlo Simulation and Time-Series Methods. *Review of Applied Science and Technology*, 3(04), 165–197. <https://doi.org/10.63125/a3047d06>
- [85]. Mokkupati, R., & Dasari, V. L. (2023). A comprehensive review on areas and applications of artificial intelligence, machine learning, deep learning, and data science. 2023 3rd international conference on innovative mechanisms for industry applications (ICIMIA),
- [86]. Mostafa, K. (2023). An Empirical Evaluation of Machine Learning Techniques for Financial Fraud Detection in Transaction-Level Data. *American Journal of Interdisciplinary Studies*, 4(04), 210-249. <https://doi.org/10.63125/60amyk26>
- [87]. Mostafa, K., & Tahmina Akter Bhuya, M. (2023). Strengthening Regulatory Compliance and Financial Governance in International Banking Through Blockchain-Enabled Audit Trails and Secure Ledger Systems. *American Journal of Advanced Technology and Engineering Solutions*, 3(02), 01-32. <https://doi.org/10.63125/e6k0e047>
- [88]. Nancy, A. A., Ravindran, D., Raj Vincent, P. D., Srinivasan, K., & Gutierrez Reina, D. (2022). Iot-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning. *Electronics*, 11(15), 2292.
- [89]. Nie, X., Fan, T., Wang, B., Li, Z., Shankar, A., & Manickam, A. (2020). Big data analytics and IoT in operation safety management in under water management. *Computer Communications*, 154, 188-196.
- [90]. Noshahri, H., olde Scholtenhuis, L. L., Doree, A. G., & Dertien, E. C. (2021). Linking sewer condition assessment methods to asset managers' data-needs. *Automation in Construction*, 131, 103878.
- [91]. Nozari, H., Szmelter-Jarosz, A., & Ghahremani-Nahr, J. (2022). Analysis of the challenges of artificial intelligence of things (AloT) for the smart supply chain (case study: FMCG industries). *Sensors*, 22(8), 2931.
- [92]. Nti, E. K., Cobbina, S. J., Attafuah, E. E., Opoku, E., & Gyan, M. A. (2022). Environmental sustainability technologies in biodiversity, energy, transportation and water management using artificial intelligence: A systematic review. *Sustainable Futures*, 4, 100068.
- [93]. Oliveira, S., Alegre, A., Carvalho, E., Mendes, P., & Proença, J. (2022). Seismic and structural health monitoring systems for large dams: theoretical, computational and practical innovations. *Bulletin of Earthquake Engineering*, 20(9), 4483-4512.
- [94]. Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education*, 20(1), 4.
- [95]. Park, J., Bai, B., Ryu, D., Liu, T., Lee, C., Luo, Y., Lee, M. J., Huang, L., Shin, J., & Zhang, Y. (2023). Artificial intelligence-enabled quantitative phase imaging methods for life sciences. *Nature Methods*, 20(11), 1645-1660.

- [96]. Pedersen, T., & Johansen, C. (2020). Behavioural artificial intelligence: an agenda for systematic empirical studies of artificial inference. *Ai & Society*, 35(3), 519-532.
- [97]. Plevris, V. (2024). AI-driven innovations in earthquake risk mitigation: a future-focused perspective. *Geosciences*, 14(9), 244.
- [98]. Plevris, V., & Papazafeiropoulos, G. (2024). AI in structural health monitoring for infrastructure maintenance and safety. *Infrastructures*, 9(12), 225.
- [99]. Popkova, E. G., & Sergi, B. S. (2024). Energy infrastructure: Investment, sustainability and AI. *Resources Policy*, 91, 104807.
- [100]. Preethichandra, D., Suntharavadivel, T., Kalutara, P., Piyathilaka, L., & Izhar, U. (2023). Influence of smart sensors on structural health monitoring systems and future asset management practices. *Sensors*, 23(19), 8279.
- [101]. Ratul, D. (2022). Engineering Resilient Flood Mitigation Using Geosynthetic and Composite Barrier Materials Performance Modeling and Environmental Impact Assessment. *Review of Applied Science and Technology*, 1(03), 100–148. <https://doi.org/10.63125/052q7d44>
- [102]. Ratul, D., & Aditya, D. (2023). AI-Driven Change Detection Using SAR, LIDAR, And Sentinel-2 Data for Landslide Monitoring and Disaster Early Warning Systems. *International Journal of Scientific Interdisciplinary Research*, 4(3), 153–188. <https://doi.org/10.63125/4y740y95>
- [103]. Ratul, D., & Subrato, S. (2022). Remote Sensing Based Integrity Assessment of Infrastructure Corridors Using Spectral Anomaly Detection and Material Degradation Signatures. *American Journal of Interdisciplinary Studies*, 3(04), 332-364. <https://doi.org/10.63125/1sdhwn89>
- [104]. Ren, Q., Li, M., Li, H., & Shen, Y. (2021). A novel deep learning prediction model for concrete dam displacements using interpretable mixed attention mechanism. *Advanced Engineering Informatics*, 50, 101407.
- [105]. Rifaai, T. M., Abokifa, A. A., & Sela, L. (2022). Integrated approach for pipe failure prediction and condition scoring in water infrastructure systems. *Reliability engineering & system safety*, 220, 108271.
- [106]. Rifat, C., & Rebeka, S. (2023). The Role of ERP-Integrated Decision Support Systems in Enhancing Efficiency and Coordination In Healthcare Logistics: A Quantitative Study. *International Journal of Scientific Interdisciplinary Research*, 4(4), 265–285. <https://doi.org/10.63125/c7srk144>
- [107]. Rigaud, M., Buekers, J., Bessems, J., Basagaña, X., Mathy, S., Nieuwenhuijsen, M., & Slama, R. (2024). The methodology of quantitative risk assessment studies. *Environmental Health*, 23(1), 13.
- [108]. Rossi, M., & Bournas, D. (2023). Structural health monitoring and management of cultural heritage structures: a state-of-the-art review. *Applied sciences*, 13(11), 6450.
- [109]. Saadane, R., Chehri, A., & Jeon, S. (2022). AI-based modeling and data-driven evaluation for smart farming-oriented big data architecture using IoT with energy harvesting capabilities. *Sustainable Energy Technologies and Assessments*, 52, 102093.
- [110]. Saghiri, A. M., Vahidipour, S. M., Jabbarpour, M. R., Sookhak, M., & Forestiero, A. (2022). A survey of artificial intelligence challenges: Analyzing the definitions, relationships, and evolutions. *Applied sciences*, 12(8), 4054.
- [111]. Sarker, I. H. (2022). AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems. *SN computer science*, 3(2), 158.
- [112]. Sarker, I. H., Hoque, M. M., Uddin, M. K., & Alsanoosy, T. (2021). Mobile data science and intelligent apps: concepts, AI-based modeling and research directions. *Mobile Networks and Applications*, 26(1), 285-303.
- [113]. Sazzadul, I., & Rebeka, S. (2024). VaR and CVaR-Based Stress Testing Using Deep Learning for Liquidity Risk Forecasting and Banking Stability Assessment. *Review of Applied Science and Technology*, 3(03), 01-30. <https://doi.org/10.63125/291phs66>
- [114]. Schmitt, M. (2023). Securing the digital world: Protecting smart infrastructures and digital industries with artificial intelligence (AI)-enabled malware and intrusion detection. *Journal of Industrial Information Integration*, 36, 100520.
- [115]. Sharma, V. B., Tewari, S., Biswas, S., Lohani, B., Dwivedi, U. D., Dwivedi, D., Sharma, A., & Jung, J. P. (2021). Recent advancements in AI-enabled smart electronics packaging for structural health monitoring. *Metals*, 11(10), 1537.
- [116]. Shehadeh, A., Alshboul, O., Al Mamlook, R. E., & Hamedat, O. (2021). Machine learning models for predicting the residual value of heavy construction equipment: An evaluation of modified decision tree, LightGBM, and XGBoost regression. *Automation in Construction*, 129, 103827.
- [117]. Shelar, A., Nile, S. H., Singh, A. V., Rothenstein, D., Bill, J., Xiao, J., Chaskar, M., Kai, G., & Patil, R. (2023). Recent advances in nano-enabled seed treatment strategies for sustainable agriculture: challenges, risk assessment, and future perspectives. *Nano-Micro Letters*, 15(1), 54.
- [118]. Siemens, G., Marmolejo-Ramos, F., Gabriel, F., Medeiros, K., Marrone, R., Joksimovic, S., & De Laat, M. (2022). Human and artificial cognition. *Computers and Education: Artificial Intelligence*, 3, 100107.
- [119]. Silionis, N. E., & Anyfantis, K. N. (2023). Data-driven probabilistic quantification and assessment of the prediction error model in damage detection applications. *Probabilistic Engineering Mechanics*, 71, 103412.
- [120]. Sivasuriyan, A., Vijayan, D. S., Devarajan, P., Stefańska, A., Dixit, S., Podlasek, A., Sitek, W., & Koda, E. (2024). Emerging trends in the integration of smart sensor technologies in structural health monitoring: A contemporary perspective. *Sensors*, 24(24), 8161.
- [121]. Sofi, A., Regita, J. J., Rane, B., & Lau, H. H. (2022). Structural health monitoring using wireless smart sensor network–An overview. *Mechanical Systems and Signal Processing*, 163, 108113.
- [122]. Spinu, N., Cronin, M. T., Enoch, S. J., Madden, J. C., & Worth, A. P. (2020). Quantitative adverse outcome pathway (qAOP) models for toxicity prediction. *Archives of Toxicology*, 94(5), 1497-1510.

- [123]. Studer, S., Bui, T. B., Drescher, C., Hanuschkin, A., Winkler, L., Peters, S., & Müller, K.-R. (2021). Towards CRISP-ML (Q): a machine learning process model with quality assurance methodology. *Machine learning and knowledge extraction*, 3(2), 392-413.
- [124]. Sun, Z., Jayasinghe, S., Sidiq, A., Shahrivar, F., Mahmoodian, M., & Setunge, S. (2024). Approach towards the development of digital twin for structural health monitoring of civil infrastructure: A comprehensive review. *Sensors*, 25(1), 59.
- [125]. Tahmina Akter Bhuya, M., & Rebeka, S. (2022). AI-Assisted Underwriting Models for Improving Risk Assessment Accuracy in U.S. Insurance Markets. *American Journal of Interdisciplinary Studies*, 3(01), 65-102. <https://doi.org/10.63125/kegg1076>
- [126]. Tansel, B., & Zhang, K. (2022). Effects of saltwater intrusion and sea level rise on aging and corrosion rates of iron pipes in water distribution and wastewater collection systems in coastal areas. *Journal of Environmental Management*, 315, 115153.
- [127]. Tasnim, K., & Anick, K. M. T. A. (2024). PLC–SCADA–Integrated Electrical Automation Frameworks for Process Optimization in Water and Wastewater Treatment Facilities. *Review of Applied Science and Technology*, 3(01), 221–262. <https://doi.org/10.63125/y1145g11>
- [128]. Tibaduiza Burgos, D. A., Gomez Vargas, R. C., Pedraza, C., Agis, D., & Pozo, F. (2020). Damage identification in structural health monitoring: A brief review from its implementation to the use of data-driven applications. *Sensors*, 20(3), 733.
- [129]. Toh, G., & Park, J. (2020). Review of vibration-based structural health monitoring using deep learning. *Applied sciences*, 10(5), 1680.
- [130]. Torero, J. L., Gerhard, J. I., Martins, M. F., Zanoni, M. A., Rashwan, T. L., & Brown, J. K. (2020). Processes defining smouldering combustion: Integrated review and synthesis. *Progress in energy and combustion science*, 81, 100869.
- [131]. Ucar, A., Karakose, M., & Kırımca, N. (2024). Artificial intelligence for predictive maintenance applications: key components, trustworthiness, and future trends. *Applied sciences*, 14(2), 898.
- [132]. Vekaria, D., Kumari, A., Tanwar, S., & Kumar, N. (2020). Şboost: an AI-based data analytics scheme for COVID-19 prediction and economy boosting. *IEEE Internet of Things Journal*, 8(21), 15977-15989.
- [133]. Viceconti, M., Pappalardo, F., Rodriguez, B., Horner, M., Bischoff, J., & Tshinanu, F. M. (2021). In silico trials: Verification, validation and uncertainty quantification of predictive models used in the regulatory evaluation of biomedical products. *Methods*, 185, 120-127.
- [134]. Wand, H., Lambert, S. A., Tamburro, C., Iacocca, M. A., O'Sullivan, J. W., Sillari, C., Kullo, I. J., Rowley, R., Dron, J. S., & Brockman, D. (2021). Improving reporting standards for polygenic scores in risk prediction studies. *Nature*, 591(7849), 211-219.
- [135]. Wang, G., & Ke, J. (2024). Literature review on the structural health monitoring (SHM) of sustainable civil infrastructure: an analysis of influencing factors in the implementation. *Buildings*, 14(2), 402.
- [136]. Wang, K., Zhao, Y., Gangadhari, R. K., & Li, Z. (2021). Analyzing the adoption challenges of the Internet of things (Iot) and artificial intelligence (ai) for smart cities in china. *Sustainability*, 13(19), 10983.
- [137]. Wu, W., Cantero-Chinchilla, S., Prescott, D., Remenyte-Prescott, R., & Chiachío, M. (2024). A general approach to assessing SHM reliability considering sensor failures based on information theory. *Reliability engineering & system safety*, 250, 110267.
- [138]. Wuni, I. Y., & Shen, G. Q. (2020). Barriers to the adoption of modular integrated construction: Systematic review and meta-analysis, integrated conceptual framework, and strategies. *Journal of Cleaner Production*, 249, 119347.
- [139]. Xiong, Z., Cui, Y., Liu, Z., Zhao, Y., Hu, M., & Hu, J. (2020). Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation. *Computational Materials Science*, 171, 109203.
- [140]. Yin, X., Chen, Y., Bouferguene, A., & Al-Hussein, M. (2020). Data-driven bi-level sewer pipe deterioration model: Design and analysis. *Automation in Construction*, 116, 103181.
- [141]. Zaheda, K., & Md Hamidur, R. (2024). GPU-Accelerated Physics-Informed Digital Twins for Real-Time State Estimation and Fault Localization in Distribution Grids. *American Journal of Scholarly Research and Innovation*, 3(02), 179-216. <https://doi.org/10.63125/msrpf04>
- [142]. Zaheda, K., & Md. Tahmid Farabe, S. (2023). Robotics and Computer Vision for Automated Inspection of Substation and Treatment-Facility Electrical Infrastructure. *Review of Applied Science and Technology*, 2(04), 194-227. <https://doi.org/10.63125/tfh15j12>
- [143]. Zhang, X., Xie, F., Ji, T., Zhu, Z., & Zheng, Y. (2021). Multi-fidelity deep neural network surrogate model for aerodynamic shape optimization. *Computer Methods in Applied Mechanics and Engineering*, 373, 113485.
- [144]. Zinno, R., Haghshenas, S. S., Guido, G., Rashvand, K., Vitale, A., & Sarhadi, A. (2022). The state of the art of artificial intelligence approaches and new technologies in structural health monitoring of bridges. *Applied sciences*, 13(1), 97.
- [145]. Zinno, R., Haghshenas, S. S., Guido, G., & Vitale, A. (2022). Artificial intelligence and structural health monitoring of bridges: A review of the state-of-the-art. *IEEE access*, 10, 88058-88078.