Waterlines 43:1 (2025), 24-33

Lifecycle Vulnerability in Urban Water Infrastructure for Predictive Maintenance Planning

Devdatta Mokashi, Sameer Sawarkar, Dr Amena Ansari, Mukund Pandurang Chougale, Gularam Masharipova, Komal Meshram

Abstract: *This paper presents a taxonomy-driven framework for lifecycle vulnerability mapping and predictive maintenance in urban water utilities under aging buried assets and budget constraints. Current planning assumes homogeneous conditions, limiting risk-aware scheduling and the integration of exposure, degradation, and failure-mode evidence. The framework integrates criticality indices, rule-based failure typologies, survival and hazard models, probabilistic degradation, and Long Short-Term Memory (LSTM) sequence models, with uncertainty quantification and feature attribution. Data sources include asset registries, network topology, telemetry, remote sensing, and socio-hydrological layers, harmonized with documented provenance and access controls. Outcomes include simulated asset-state trajectories, annual failure probabilities, risk hotspot maps transformed into ranked interventions, and schedule optimization under crew and budget constraints; model quality is assessed with Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and coefficient of determination (R2) using temporal splits, back casting, and cross-validation with 95 percent confidence bounds. Sensitivity analyses examine budget limits, model variants, and nonstationary climate and land-use scenarios. Because many results are simulation-derived and labels are sparse, effects are reported qualitatively and utility-specific magnitudes may vary. The*

Devdatta Mokashi, Bharati Vidyapeeth's college of Engineering Lavale Pune, Maharashtra, India. dev.mokashi@gmail.com
Sameer Sawarkar, Department of Civil Engineering, Pimpri Chinchwad College of Engineering & Research Ravet, Pimpri-Chinchwad, Maharashtra 412101, India. sameer.sawarkar@pccoer.in
Dr Amena Ansari, Department of Civil Engineering, Deogiri Institute of Engineering and Management Studies,Deogiri Campus, Aurangabad, Maharashtra 431005, India. amenatamboli2011@gmail.com
Mukund Pandurang Chougale, Department of Civil Engineering, D Y Patil College of Engineering, Akurdi, Pune-411044, India. mpchougale@gmail.com
Gularam Masharipova, Department of philosophy, Alfraganus University, Tashkent, Uzbekistan masharipovagularam2006@gmail.com
Komal Meshram, Department of Civil Engineering, Tulsiramji Gaikwad Patil College of Engineering and Technology, Nagpur, Maharashtra 441108, India. komalmeshram941@gmail.com

contribution is a unified, auditable workflow that links forecasting, risk, and cost into actionable schedules with transparent uncertainty. Practically, the framework enables planners and policymakers to prioritize interventions that mitigate failure risk while smoothing lifecycle costs under real-world constraints.

Keywords: Lifecycle Planning, Vulnerability Mapping, Predictive Maintenance, Asset Management, Urban Water Infrastructure, Failure Prevention

Introduction

Urban water utilities face aging buried assets, tight budgets, and cascading service and health risks from failures. Although stressors vary by land use, hydrology, climate, and capacity, planning assumes homogeneity; risk-aware scheduling suffers. We target gaps in linking lifecycle theory to maintenance timing, mapping vulnerability points by asset and failure mode, and using public degradation data. We propose a taxonomy-driven framework integrating criticality and planning heuristics to simulate states, prioritize risk-aware interventions, and quantify cost disruption trade-offs, with transfer suggested by scenario-driven runoff modelling (Subbarayan et al., 2025) and machine-learning vulnerability assessment (Mondal et al., 2025) as hypotheses.

Literature Review

Table 1. Representative studies by domain, method, data source, and key metric

Study	*Domain*	*Method*	*Data source*	*Key metric*
Yin et al., 2025	Wastewater shock-load prediction	Probabilistic encoder-decoder LSTM	Real WWTP time series	Accuracy improvement 49.7 percent, over-limit coverage to 6 h
Makhlouf et al., 2025	Groundwater quality monitoring	Gaussian Process Regression and other ML	246 samples with EC, pH, site conditions	Correlation coefficient up to 0.97

Karim et al., 2025	Urban expansion exposure mapping	CA Markov chain with GIS	Landsat imagery 1998-2023	Urbanized area 53.6 percent 2023, 75.8 percent 2048

This table (1) summarizes exemplar studies and their domains, methods, data sources, and key reported metrics.

This review synthesizes methods for lifecycle vulnerability mapping and predictive maintenance; although families target distinct problems, they inform risk triage and scheduling. Probabilistic event models improve shock-load prediction (Yin et al., 2025), ML monitoring pipelines streamline water-quality assessment (Makhlouf et al., 2025), and geospatial urban-expansion analyses map exposure (Karim et al., 2025). Studies should report calibration, uncertainty, RMSE and R2 with units, and cost and risk. Asset registers, SCADA, forecasts, and remote-sensing enable inference, are patchy, and claims on transferability, robustness, and sparsity sensitivity require empirical support. Evidence remains thin on exposure-degradation integration and operational shock-load deployments. Units anchor interpretation.

Materials and Methods

Although datasets vary by utility, the section must specify provenance and attributes for asset registries, network topology, material and install year, failures and maintenance, hydraulic loading, and socioeconomic or consequence indicators; define quality metrics, missingness treatment, harmonization, and georeferencing. Justify models (survival or hazard, probabilistic degradation with uncertainty, rule-based failure typologies) against physics and data. Report uncertainty probabilistically, validate via cross-validation or back casting, run scenarios and sensitivity. Translate risk (likelihood, consequence, cost) into priorities, optimize schedules under constraints, simulate strategy variants, quantify risk and lifecycle-cost impacts, document assumptions, versioning, integration, access controls, equity, privacy, and flag assumptions for external validation.

Data Sources

This section details datasets for lifecycle vulnerability mapping. Although registers and failure logs differ by licensing and completeness, we record provenance, ownership, coverage, and resolution for assets, material/age, sensors, remote-sensing LULC, and socio-hydrological layers; outputs carry uncertainty (e.g., change-detection intervals, missing-log bias). Georeferencing uses base maps, harmonization via controlled vocabularies, and mismatches use spatial joins, temporal alignment. Urban expansion and LULC projections from GEE/ML and QGIS-MOLUSCE inform exposure through buffer overlaps (Karim et al., 2025; Gunduz, 2025). Preprocessing documents CRS, filtering, gap-filling, with calibration and holdouts. Access constraints and records limit generalization; machine-readable metadata, versioning, and conservative inference are mandatory.

Model Design

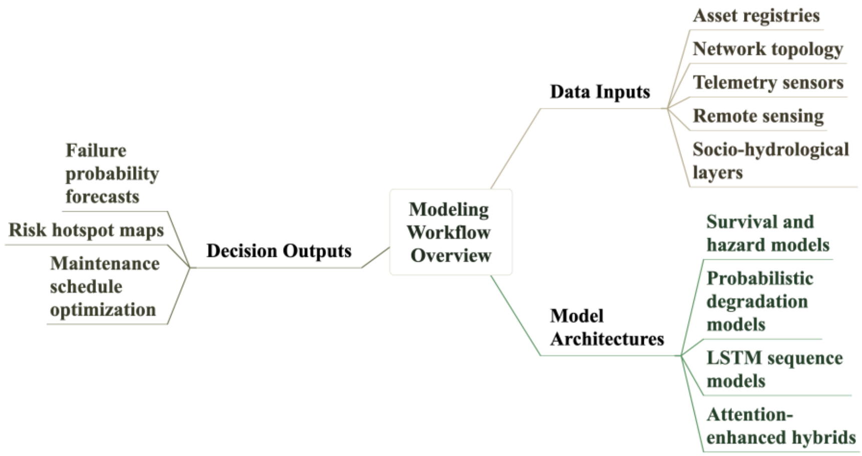

Figure 1. Proposed modeling workflow overview

This figure (1) summarizes data inputs, model architectures, and decision outputs linking lifecycle vulnerability mapping to predictive maintenance scheduling.

Although urban water assets fail heterogeneously, we predict failure probability and risk to schedule and balance cost-risk trade-offs. Inputs span attributes, telemetry, environment, failures, and degradation proxies, missingness addressed via imputation and temporal smoothing, with typologies embedded and spatiotemporal lags. Temporal behaviour employs survival or point-process models plus ensemble LSTMs for robustness (Hosseini et al., 2025), while attention-enhanced LSTM-Transformer hybrids capture dependence (Ren et al., 2025). Decision requires uncertainty quantification and feature attribution for stakeholder acceptance. Simulations test applicability, risk reduction, cost impact, scheduling under budgets. Limitations include scarce labels, covariate shift, and computational scaling—claims remain bounded by cited evidence.

Evaluation Metrics

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (1)$$

Equation (1) defines RMSE as the square-root average of squared residuals for scale-sensitive comparison and highlights sensitivity to large errors in evaluation.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \widehat{y_i}| \tag{2}$$

Equation (2) defines MAE as the average absolute residual magnitude to provide a scale-consistent and more outlier-robust error summary than squared-loss metrics.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{3}$$

Equation (3) states the coefficient of determination as the variance-explained fraction, emphasizing interpretive limits under nonlinearity or heteroskedastic residuals.

Although RMSE, MAE, and R^2 are standard, their use for asset-failure prediction demands explicit justification (Ren et al., 2025). RMSE amplifies large residuals, MAE offers outlier-robust errors, and R^2 quantifies variance explained but can mislead under nonlinearity or heteroskedasticity (Ren et al., 2025). Rare, high-impact failures require tail-aware diagnostics and evaluation that preserves event characteristics; temporal splits, extreme-focused validation, and uncertainty quantification (Hiraga & Meza, 2025). Nonstationary climate and land-use reshape error distributions and baselines, so distribution-aware checks and cost-weighted errors aligned to service disruption are needed (Liu et al., 2025). Cross-domain transfers must state risk assumptions (Ren et al., 2025).

Comparative Analysis

We specify rigorous, objective benchmarking for shared urban water and hydrological datasets. Although single-catchment studies inform practice, claims require qualification and multi-period tests. Standardize units, lead-time, and metric computation; report Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R2 with temporal transferability, sensitivity to nonstationarity, computational cost, data needs, and interpretability. Climate-forecast forcings change skill by lead time (Girons Lopez et al., 2025). Catchment-wise ensemble LSTMs improve accuracy (Hosseini et al., 2025). Probabilistic deep learning enhances shock handling and uncertainty quantification (Yin et al., 2025). Mandate data splits, cross-validation, lead-time windows, calibration, and uncertainty propagation for maintenance scheduling.

Results

Outputs included asset-state trajectories, hazard functions, lifecycle costs by schedule, and intervention risk reduction, plus optimization efficacy, adaptability, and applicability. Although many outputs are simulation-derived, we validated failure timing and rates against operational records via withheld periods and reported cross-validated errors with 95 percent confidence bounds. Sensitivity covered budget limits, degradation-model variants, and climate loadings; parameter elasticities and interval widths guided prioritization and robustness. Costs dominate in sparse districts. Remote proxies were applied only where super-resolution sharpened detection (Shu et al., 2025), and climate attributions were limited to patterns consistent with semiarid lake responses (Jimenez-Bonilla et al., 2025).

Risk Reduction

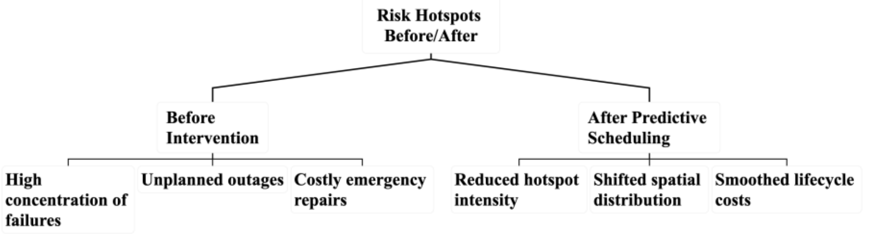

Figure 2. Spatial risk hotspots before after

This figure (2) contrasts mapped risk concentrations before and after predictive-maintenance scheduling, illustrating shifts in hotspot intensity and location.

Operational risk reduction follows when hotspot maps become ranked interventions tied to expected failure frequency, spatial concentration, maintenance cost impact, and schedule optimization efficacy. Although maps suggest where, decisions require outputs such as annual failure probabilities, expected failures, and constrained schedules with intervals. Optimized replacement or refurbishment from degradation models and asset data can shift risk and reallocate budgets; any percentage change or cost-benefit must be empirically shown. Climate exposure or hybrid ML effects need validation (Mondal et al., 2025). Lab analogies motivate, not generalize, so field trials are needed (Hassanpour et al., 2025). Align with budgets, regulation, procurement.

Cost Impact

This section requires converting vulnerability maps into monetary metrics. Although risk scores guide priorities, decisions must rest on NPV, avoided outages and emergency repair costs, capex smoothing, and risk-weighted losses. Analyses must quantify engineering and societal costs, disclose inputs and assumptions, and trace estimates to sources. Uncertainty warrants sensitivity tests and probabilistic scenarios; report limitations and robustness checks. Compare replacement, condition-based renewal, and deferred repair, and justify hydro-mechanical or lab-derived assumptions with simulations and scale-up caveats (Xiao et al., 2025; Hassanpour et al., 2025).

Schedule Optimization

This section defines requirements for predictive maintenance scheduling. Although data richness varies, link degradation trajectories to constrained scheduling to minimize lifecycle risk and cost while maintaining service. Make uncertainty in failure timing, consequence severity, and parameters explicit; quantify gains via probabilistic or scenario simulations and embed criticality and failure-mode classes in objectives that value heterogeneous consequences. Translate vulnerability maps into time-to-failure distributions, stress-conditional failure, and crew constraints. Use sensitivity analysis for calibration and multi-resolution BIM hydro-mechanical coupling to refine deterioration and evidence budget-realistic schedule shifts; evaluate metrics and provide planner outputs (Kheshti Azar et al., 2025; Xiao et al., 2025).

Discussion

Policy Implications

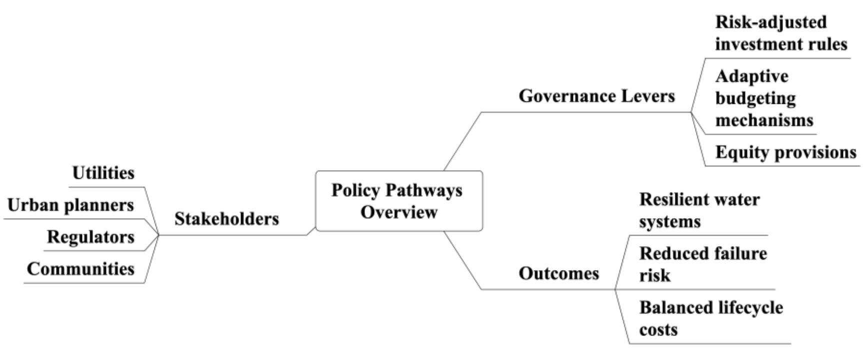

Figure 3. Policy levers for resilient water systems

This figure (3) maps governance levers to outcomes and stakeholders

Although models pinpoint hotspots, capital allocation must apply risk-adjusted rules weighting failure likelihood, criticality, lifecycle cost, and service continuity. Adaptive budgeting and waivers or incentives expedite priority actions with trade-offs recorded. Execution requires coordination across planning, utilities, land-use, supported by data-sharing protocols and metadata standards. Land-use and climate projections should steer anticipatory scheduling to avoid maladaptation, reflecting urban expansion patterns, hybrid ML amplification, and CA-Markov GIS demand forecasts (Karim et al., 2025; Mondal et al., 2025; Tahir et al., 2025). Equity provisions must limit disproportionate disruption. Performance metrics anchor oversight—applicability, risk reduction, cost impact, scheduling efficacy, adaptability—tested via scenario-based simulations.

Limitations

This section outlines limitations on lifecycle vulnerability mapping and predictive maintenance. Although public registries are incomplete, failure logs biased, monitoring coarse, and covariates misaligned with buried-asset scales, inference should be qualified. Gains from higher-resolution environmental inputs require evidence from super-resolution studies (Shu et al., 2025). Models assume parametric degradation, fixed failure typologies, and calibration on few events; these choices risk overfitting and weak generalization. Uncertainty quantification and sensitivity analysis are required, not deterministic forecasts. External drivers require GCM-based streamflow evidence for hydrological loading (Girons Lopez et al., 2025). Transferability hinges on practice, budgets, and governance, unvalidated claims are provisional.

Conclusion

This paper advances predictive maintenance by unifying degradation models, failure-mode typologies, and scheduling heuristics into lifecycle vulnerability mapping. Although data and model assumptions constrain transferability, the scheme enables resource-constrained, risk-aware prioritization; cost-effectiveness or schedule gains need audited pilots or counterfactual simulations estimating cost per failure avoided and schedule adherence. Sensitivity to climate extremes and land-use change demands empirical backing and rigorous scenario design (Liu et al., 2025; Hiraga & Meza, 2025). Next steps include integration into asset-management workflows and stakeholder engagement to align outputs with regulation. Priorities include calibration, extreme-event scenarios, metrics for realized risk reduction, and transparent uncertainty communication.

References

Subbarayan S.; Youssef Y.M.; Singh L.; Dąbrowska D.; Alarifi N.; Ramsankaran R.A.A.J.; Visweshwaran R.; Saqr A.M. (2025). Soil and Water Assessment Tool-Based Prediction of Runoff Under Scenarios of Land Use/Land Cover and Climate Change Across Indian Agro-Climatic Zones: Implications for Sustainable Development Goals. *Water (Switzerland)*, 17(3). DOI: 10.3390/w17030458.

Girons Lopez M.; Bosshard T.; Crochemore L.; Pechlivanidis I.G. (2025). Leveraging GCM-based forecasts for enhanced seasonal streamflow prediction in diverse hydrological regimes. *Journal of Hydrology*, 650. DOI: 10.1016/j.jhydrol.2024.132504.

Yin H.; Chen Y.; Zhou J.; Xie Y.; Wei Q.; Xu Z. (2025). A probabilistic deep learning approach to enhance the prediction of wastewater treatment plant effluent quality under shocking load events. *Water Research X*, 26. DOI: 10.1016/j.wroa.2024.100291.

Makhlouf A.; El-Rawy M.; Kanae S.; Sharaan M.; Nada A.; Ibrahim M.G. (2025). Streamlining the monitoring and assessment of irrigation groundwater quality using machine learning techniques. *Environmental Earth Sciences*, 84(5). DOI: 10.1007/s12665-025-12132-4.

Hiraga Y.; Meza J. (2025). Extreme precipitation modeling and Probable Maximum Precipitation (PMP) estimation in Chile. *Journal of Hydrology: Regional Studies*, 58. DOI: 10.1016/j.ejrh.2025.102274.

Karim S.; Chen Y.; Kayumba P.M.; Ahmad I.; Iqbal H. (2025). Comprehensive geospatial analysis of urban expansion dynamic in Lahore, Pakistan (1998–2023). *Journal of Urban Management*, 14(2), pp. 577. DOI: 10.1016/j.jum.2024.11.012.

Xiao F.; Chen X.; Zhu Y.; Xie P.; Salimzadeh S.; Zhang Q.-B. (2025). Multi-LoD BIM integrated design framework for pressurised tunnel: Hydro-mechanical coupling simulation and sustainability assessment. *Tunnelling and Underground Space Technology*, 158. DOI: 10.1016/j.tust.2025.106404.

Hosseini F.; Prieto C.; Álvarez C. (2025). Ensemble learning of catchment-wise optimized LSTMs enhances regional rainfall-runoff modelling − case Study: Basque Country, Spain. *Journal of Hydrology*, 646. DOI: 10.1016/j.jhydrol.2024.132269.

Kheshti Azar M.; Giudicianni C.; Creaco E. (2025). Sensitivity Analysis-Aided Calibration of Urban Drainage Modeling. *Water (Switzerland)*, 17(5). DOI: 10.3390/w17050612.

Ren D.; Hu Q.; Zhang T. (2025). EKLT: Kolmogorov-Arnold attention-driven LSTM with Transformer model for river water level prediction. *Journal of Hydrology*, 649. DOI: 10.1016/j.jhydrol.2024.132430.

Huang X.; Wang H.; Song X.; Han Z.; Shu Y.; Wu J.; Luo X.; Zheng X.; Fan Z. (2025). Ecological risks of PFAS in China's surface water: A machine learning approach. *Environment International*, 196. DOI: 10.1016/j.envint.2025.109290.

Yu H.; Jiang J.; Gu X.; Cao C.; Shen C. (2025). Using dynamic Bayesian belief networks to infer the effects of climate change and human activities on changes in regional ecosystem services. *Ecological Indicators*, 170. DOI: 10.1016/j.ecolind.2024.113023.

Liu J.; Li M.; Li R.; Shalamzari M.J.; Ren Y.; Silakhori E. (2025). Comprehensive Assessment of Drought Susceptibility Using Predictive Modeling, Climate Change Projections, and Land Use Dynamics for Sustainable Management. *Land*, 14(2). DOI: 10.3390/land14020337.

Jiménez-Bonilla A.; Rodríguez-Rodríguez M.; Yanes J.L.; Gázquez F. (2025). Impact of climate change on permanent lakes in a semiarid region: Southwestern Mediterranean basin (S Spain). *Science of the Total Environment*, 961. DOI: 10.1016/j.scitotenv.2024.178305.

Tahir Z.; Haseeb M.; Mahmood S.A.; Batool S.; Abdullah-Al-Wadud M.; Ullah S.; Tariq A. (2025). Predicting land use and land cover changes for sustainable land management using CA-Markov modelling and GIS techniques. *Scientific Reports*, 15(1). DOI: 10.1038/s41598-025-87796-w.

Mondal I.; Das A.; Hossain S.A.; Jose F.; Altuwaijri H.A. (2025). Climate change-induced vulnerability assessment for the Florida Coast using hybrid machine learning models. *Ecological Indicators*, 171. DOI: 10.1016/j.ecolind.2025.113242.

Gündüz H.İ. (2025). Land-Use Land-Cover Dynamics and Future Projections Using GEE, ML, and QGIS-MOLUSCE: A Case Study in Manisa. *Sustainability (Switzerland)*, 17(4). DOI: 10.3390/su17041363.

Shu P.; Aslam R.W.; Naz I.; Ghaffar B.; Kucher D.E.; Quddoos A.; Raza D.; Abdullah-Al-Wadud M.; Zulqarnain R.M. (2025). Deep Learning-Based Super-Resolution of Remote Sensing Images for Enhanced Groundwater Quality Assessment and Environmental Monitoring in Urban Areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18, pp. 7933. DOI: 10.1109/JSTARS.2025.3548010.

Aldabagh H.A.; Ibrahim R.T. (2025). Hybrid Intelligent Technique between Supervised and Unsupervised Machine Learning to Predict Water Quality. *International Journal of Computing and Digital Systems*, 17(1). DOI: 10.12785/ijcds/1571031447.

Hassanpour N.; Lara J.L.; Vicinanza D.; Contestabile P. (2025). Laboratory experiments on transforming emerged into submerged breakwaters. *Ocean Engineering*, 324. DOI: 10.1016/j.oceaneng.2025.120610.