Joint CSCE Construction Specialty & CRC Conference 2025
Conférence conjointe spécialisée en construction de la SCGC et CRC-2025

Montreal, Quebec
July 28-31, 2025 / *28-31 juillet 2025*

Risk-Based Prioritization of Water Main Replacements under Climate Change Scenarios

Khashei, M.[1], Boloukasli, F.[1], Hedaiaty Marzouny, N.[1], and Dziedzic, R.[1]

[1] Dept of Building, Civil & Environmental Engineering, Concordia University, Montréal, QC, Canada

ABSTRACT: Ensuring the resilience of water distribution networks is essential for sustainable urban development, particularly in the face of climate change and aging infrastructure. However, many parts of these systems in North America have exceeded their expected service life, leading to an increasing frequency of failures. This study presents a comprehensive risk-based framework for prioritizing water main replacements by integrating machine learning-based probability of failure (PoF) assessments with consequence of failure (CoF) evaluations. The methodology leverages historical failure records, environmental conditions, and climate projections under three Shared Socioeconomic Pathway (SSP) scenarios. The City of Waterloo, Ontario, serves as a case study, providing real-world validation of the proposed approach. To estimate PoF, four machine learning models—Random Forest, k-Nearest Neighbors, Artificial Neural Networks, and Light Gradient Boosting Machine—were tested, with LightGBM demonstrating the highest F1-score. The CoF assessment incorporates economic, social, and environmental dimensions, ensuring a holistic evaluation of failure consequences. Results indicate that climate change significantly influences failure risk, with extreme temperature fluctuations accelerating pipeline deterioration. Warmer climate scenarios (SSP2 and SSP5) lead to a greater proportion of high-risk water mains, necessitating proactive infrastructure management strategies. By integrating advanced predictive modeling with a comprehensive risk assessment approach, this study provides municipalities with a data-driven decision-making tool for infrastructure renewal. The findings highlight the importance of climate-adaptive planning to enhance the resilience and sustainability of water distribution networks amid evolving environmental challenges.

1. INTRODUCTION

Access to a safe and reliable water supply is essential for both public health and economic stability (Gleick, 2002). Water utilities play a crucial role in ensuring the continuous availability of clean water, both in terms of quality and quantity, to meet the needs of communities. Pipe infrastructure is fundamental to maintaining this supply, safeguarding water security and quality throughout the distribution network. While advancements in science, technology, and financial resources have improved drinking water safety in developed regions, water quality-related incidents still occur due to technical failures, institutional lapses, and, in some cases, managerial negligence (Hrudey & Hrudey, 2004). To address these risks effectively, a structured approach is essential, with risk-based decision-making offering a methodical framework to recognize potential hazards, evaluate their impacts, and implement proactive measures to enhance resilience and reliability in water utilities (MacGillivray & Pollard, 2008).

Risk-based assessment in water infrastructure management involves three stages: identifying potential failure scenarios, assessing the failure likelihood, and evaluating failure consequences. This process guides the prioritization of rehabilitation and replacement in pipe networks, focusing on the most critical risks (Mamo, 2015). Several studies have identified factors contributing to water main pipeline failures, encompassing physical, environmental, and operational components (Hussein Farh et al., 2023). Almheiri et al. (2023) examined a wide range of factors affecting water pipeline failures and found that climatic conditions such as air temperature, precipitation, and evaporation play a significant role, alongside physical factors like pipe material and size. Climate change significantly affects water infrastructure, influencing the frequency of failures in water distribution networks (Fan et al., 2023; Khashei, Boloukasli ahmadgourabi, et al., 2024). Rising temperatures, prolonged droughts, and extreme weather events contribute to soil settlement, pipe material degradation, and increased stress on aging infrastructure, ultimately compromising the reliability of water systems (Ahmad et al., 2023). These findings highlight the importance of assessing climate change scenarios to better inform infrastructure management strategies.

Advanced predictive modeling techniques have demonstrated significant potential in addressing water infrastructure challenges, including precise failure forecasting and improved system performance (Boloukasli ahmadgourabi & Dziedzic, 2024). As the next stage in risk-based water infrastructure management, incorporating these models into risk-based assessment frameworks strengthens proactive maintenance strategies and improves resource allocation efficiency. Barton et al. (2022) categorize failure prediction models into three main categories: deterministic, probabilistic, and machine learning. Among these, machine learning models have gained significant attention due to their ability to process large, complex datasets and capture nonlinear relationships. Recently, researchers have increasingly turned to machine learning techniques to address the challenge of pipe failures in urban water distribution systems. Commonly used models include Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Evolutionary Polynomial Regression (EPR), and tree-based models, all of which have demonstrated strong predictive performance and contributed to enhancing the resilience of water infrastructure (Fan et al., 2022; Latifi et al., 2024).

Evaluating failure consequences represents the final stage of risk-based water infrastructure management, focusing on the potential impacts of system failures on communities and services. Previous studies have primarily concentrated on hydraulic reliability aspects and the associated economic costs of service disruptions. For instance, Creaco et al. (2012) evaluated network reliability through pressure-driven hydraulic models to assess demand shortfalls, while Kanta & Brumbelow (2013) examined the risks of diminished water distribution performance during critical events like urban fires. However, there is a growing recognition of the need for a more comprehensive framework that integrates not only economic and hydraulic factors but also social and environmental considerations. This broader approach ensures that water infrastructure management strategies are resilient and equitable, addressing the diverse consequences of failures in an increasingly complex and interconnected urban environment.

The novelty of this paper lies in two key contributions to risk-based water main replacement prioritization. First, the study incorporates climate change effects into the initial two stages of risk-based assessment—identifying failure scenarios and assessing failure likelihood. By integrating climate-related factors such as temperature extremes, increased precipitation variability, and shifts in soil conditions, the model captures the potential impacts of changing environmental conditions on pipeline deterioration and failure probability. This approach provides a dynamic and proactive assessment of infrastructure vulnerability under evolving climate scenarios. Second, the paper introduces a comprehensive framework for evaluating the consequences of failures, extending beyond traditional metrics like hydraulic reliability and cost. This framework integrates economic, social, and environmental dimensions, including factors such as replacement costs, traffic delays, land use, and the number of residents affected. By considering these broader impacts, the approach ensures a holistic understanding of failure consequences on both communities and ecosystems.

The remainder of this paper is structured as follows: Section 2 outlines the methodology, Section 3 presents the case study and results, and Section 4 discusses the implications and concludes the study.

2. MATERIALS AND METHODS

The methodology of this study establishes a risk-based framework for prioritizing water main replacements under climate change scenarios. This framework integrates both the Probability of Failure (PoF) and Consequence of Failure (CoF) to support data-driven decision-making in infrastructure renewal. PoF is estimated using machine learning models trained on a dataset encompassing pipe attributes, historical failures, road networks, and climate data to capture environmental stressors affecting pipeline deterioration. CoF is assessed by incorporating economic, social, and environmental factors to quantify the broader impacts of pipe failures. The following subsections outline the case study (2.1), data preprocessing (2.2), failure consequence assessment (2.3), failure probability modeling (2.4), and risk assessment (2.5).

2.1 Case Study

For this research, the City of Waterloo, Ontario, is selected as the case study due to its aging water infrastructure and rapid urban growth. The city's 400 km water distribution network, primarily sourced from groundwater, faces increasing demand, with a 15.7% population growth from 2016 to 2021 and projections of 140,500 residents by 2041. The Asset Management Plan (AMP) rates the current system as "good", with expected deterioration to "fair" in 25 years, highlighting the need for proactive rehabilitation (City of Waterloo., 2024). This study utilizes two datasets: a water network inventory covering 7,692 assets with attributes such as material, diameter, installation year, lining type, and pipe length, and a watermain break dataset (City of Kitchener, 2024). Additionally, since this research examines the influence of climate change on water main breaks, historical weather data— including minimum, maximum, and mean temperatures, as well as precipitation levels—has been collected from Environment and Climate Change Canada (ECCC) for analysis (Environment and Climate Change Canada, 2024).

2.2 Data Preprocessing

To ensure dataset reliability, missing values, inconsistencies, and outliers were removed, and categorical variables were encoded where necessary. The cleaned datasets—including water main inventory, break records, and climate data—were then merged using a unique pipe identifier and corresponding time periods for a unified analysis. Specifically, the water main inventory provided key physical attributes such as Diameter, Material, Length, Lining Status, and Age, while the break records contributed failure-related variables including Historical Failure Count (representing the number of past failures), Cumulative Failure Duration (CFD), Cumulative Time Downtime (CTD), and Cumulative Hydraulic Damage (CHD).

Given the study's objective of predicting future water main failures, decade-long intervals were used instead of annual records to address data imbalance and improve modeling efficiency. This approach allows for a more effective assessment of cumulative failures, pipe age, and climate-related factors over time. A total of five 10-year intervals were created from 1975 to 2024. For evaluating the performance of machine learning models, the most recent interval (2015–2024) was reserved for testing, while the remaining data was used for training.

Additionally, climate-related variables were incorporated to evaluate the long-term effects of climate variability on pipeline deterioration. These variables—including mean temperature, intensities of air temperature changes, daily temperature variation, freezing and thawing days and indices, and rainfall indices—were explicitly included as input features in the machine learning models. This ensured that the predictions directly captured the impacts of climate change and environmental conditions on pipeline deterioration and probability of failure (Khashei, Dziedzic, et al., 2024). allowing for a comprehensive assessment of environmental impacts on water main failures. To account for future climate variability, projections of temperature and precipitation were obtained from Environment and Climate Change Canada simulations for three Shared Socioeconomic Pathway (SSP) scenarios: SSP1, SSP2, and SSP5. SSP1 represents a low-emission, sustainable future, potentially limiting warming to below 2°C. SSP2 follows a moderate pathway, characterized by uneven development and an emissions peak in the mid-century, leading to moderate warming. In contrast, SSP5 describes a high-emission, fossil-fuel-dependent scenario, resulting in significant global warming (Eyring et al., 2016). The city of Waterloo's downscaled future weather data, aligned with these scenarios, was extracted for further analysis.

2.3 Probability of Failure (PoF)

To estimate the Probability of Failure (PoF) for water mains, four machine learning (ML) models were implemented and compared: Random Forest (RF), K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), and Light Gradient-Boosting Machine (LightGBM). These models were selected due to their established effectiveness in predicting pipeline failures based on historical break records and environmental conditions.

Random Forest (RF) is an ensemble learning algorithm that constructs multiple decision trees and aggregates their predictions to enhance accuracy and reduce overfitting (Breiman, 2001). LightGBM is a fast, efficient, and high-performance gradient boosting framework designed for large-scale data. It uses leaf-wise growth, reducing training time and improving accuracy, making it well-suited for classification and regression tasks, especially with imbalanced datasets (Ke et al., 2017). Preliminary testing also included comparisons with XGBoost, another widely used gradient boosting model. However, LightGBM was selected due to its superior balance between precision and recall (highest F1-score) and faster training times compared to XGBoost, providing computational advantages for large-scale predictive modeling tasks (Alshari et al., 2021). K-Nearest Neighbors (KNN) is a distance-based classification algorithm that assigns a new data point to the most common class among its nearest neighbors, assuming that similar instances exist in close proximity (Cover & Hart, 1967). Artificial Neural Networks (ANN), inspired by biological neural systems, consist of interconnected layers of neurons that learn complex patterns through backpropagation, with the model's architecture optimized through a trial-and-error process (Rumelhart et al., 1986). The hyperparameters of each model were optimized using Randomized Search Cross-Validation to enhance predictive performance, ensuring the models were fine-tuned for this application.

2.4 Consequence of Failure (CoF)

The Consequence of Failure (CoF) is evaluated based on economic, environmental, and social factors to assess the potential impact of water main failures. This analysis incorporates attributes from the water main inventory dataset, including pipe material, size, and installation year, along with external datasets related to road classifications, proximity to water bodies, environmentally sensitive areas, and parks. These datasets were spatially linked using a 100-meter proximity criterion in GIS, allowing for a comprehensive assessment of failure consequences.

Economic Considerations
The economic impact of failure is primarily influenced by road classification, pipe material, and pipe size. Water mains located beneath high-traffic roads pose a greater risk of economic disruption due to potential traffic delays and repair complexities. Additionally, pipe material affects maintenance costs, as some materials deteriorate faster or are more expensive to replace. Pipe size is also critical; larger pipes typically serve broader areas, and their failure can result in widespread service disruptions and higher financial losses. Economic factors were assessed using the water distribution inventory from the City of Waterloo, which includes data on pipe material, pipe size, and road classification (City of Kitchener, 2024).

Environmental Considerations
Failures near water bodies increase the risk of contamination, while those in environmentally sensitive areas can cause ecological harm, making these locations a priority in risk assessments. Larger pipes also pose a greater environmental threat, as their failure can lead to significant water loss and potential landscape degradation (Folkman, 2018). Environmental risk was evaluated using pipe size data from the water distribution inventory of the City of Waterloo, combined with GIS-based spatial analysis. The proximity of linear watermain features to water bodies and environmentally sensitive areas (polygon layers) was assessed using spatial distance analysis in GIS (City of Kitchener, 2024).

Social Considerations

From a social perspective, road classification is a key factor, as failures in densely populated areas or near critical infrastructure—such as hospitals or schools—can lead to significant disruptions. Additionally, water mains near parks and recreational areas are prioritized to prevent interruptions in water supply that could impact public health and community well-being. Additionally, pipe size is relevant, as failures in larger pipes typically affect a greater number of residents, amplifying social consequences. Social factors were evaluated by incorporating pipe size and road classification data from the water distribution inventory of the City of Waterloo, alongside GIS-based proximity analysis. Spatial analysis was used to determine the distance between linear watermain features and parks (polygon layer) to assess their potential social impact (City of Kitchener, 2024).

To quantify the Consequence of Failure, a structured framework was developed to assign scores to each contributing factor within the economic, environmental, and social perspectives. A weighted sum of these scores was then calculated for each perspective, ensuring a balanced assessment of failure impacts. For the final CoF estimation, different weights were assigned to the economic, environmental, and social components to reflect their relative significance in infrastructure management. The framework and weighting criteria were determined through a combination of literature review and expert consultation, ensuring that both empirical evidence and practical insights informed the assessment methodology. This approach captures the diverse impacts of water main failures while maintaining consistency with industry best practices. It is assumed that CoF values remain constant over the study period, as the fundamental risks associated with failure locations, land use, and infrastructure characteristics are not expected to change significantly during this timeframe.

2.5 Risk Assessment

The risk associated with each water main was evaluated by combining the Probability of Failure (PoF) and Consequence of Failure (CoF). The risk score for each pipe was computed as the product of these two factors, providing a quantitative measure of failure risk (Rahman et al., 2014). Based on the calculated scores, the risk levels were categorized into three groups: Low (0–15), Medium (15–30), and High (30–100). Pipes classified as High-risk should undergo immediate detailed inspection and be prioritized for rehabilitation or replacement within short-term planning horizons. Medium-risk pipes are scheduled for regular monitoring and proactive inspection, allowing timely intervention if conditions worsen. Low-risk pipes are managed through routine monitoring, with maintenance or replacement planned as part of long-term infrastructure renewal schedules. These thresholds were determined based on historical maintenance data and insights from relevant literature, ensuring efficient resource allocation and preventing unnecessary replacements (Godfrey et al., 2011). This helps balance proactive intervention with the risk of over-replacement, ensuring cost-effectiveness in long-term asset management. This risk-based approach enhances decision-making by allowing municipalities to allocate resources efficiently, addressing the most critical vulnerabilities within the water distribution system.

3. RESULTS AND DISCUSSION

Table 1 presents the performance of different ML models for predicting the PoF of water mains. While all models show high accuracy, this metric is misleading due to the dataset's imbalance—only 3% of pipes have experienced failure. Therefore, recall, precision, and F1-score are more relevant for evaluating model performance. Random Forest (RF) achieves the highest recall (0.718), detecting more failures but with low precision (0.331), leading to frequent false positives. k-Nearest Neighbors (kNN) has the highest precision (0.654) but the lowest recall (0.385), missing many failures. Artificial Neural Networks (ANN) capture a reasonable portion of failures but still suffers from low precision (0.423). LightGBM, on the other hand, achieved a balanced trade-off, with a recall of 0.542 and precision of 0.516, leading to the highest F1-score (0.528) among all models. This suggests that LightGBM offers a more effective balance between capturing failures and minimizing false positives. Given its superior performance in terms of F1-score—an essential metric for imbalanced classification problems— LightGBM is selected for future predictions under three climate change scenarios and subsequent risk calculations.

Table 1: Performance of each ML model.

Model	Accuracy	Recall	Precision	F1-Score
ANN	0.938	0.605	0.423	0.497
kNN	0.959	0.385	0.654	0.485
LightGBM	0.951	0.542	0.516	0.528
RF	0.913	0.718	0.331	0.453

The risk of failure framework encompasses both the consequence and probability of failure. CoF evaluates the impact of failure across three aspects: Economic, Environmental, and Social. PoF is determined entirely by the probability of failure of the pipeline, where a higher failure probability corresponds to a higher score. Figure 1 illustrates this framework, which helps prioritize pipelines for maintenance and replacement based on risk levels.

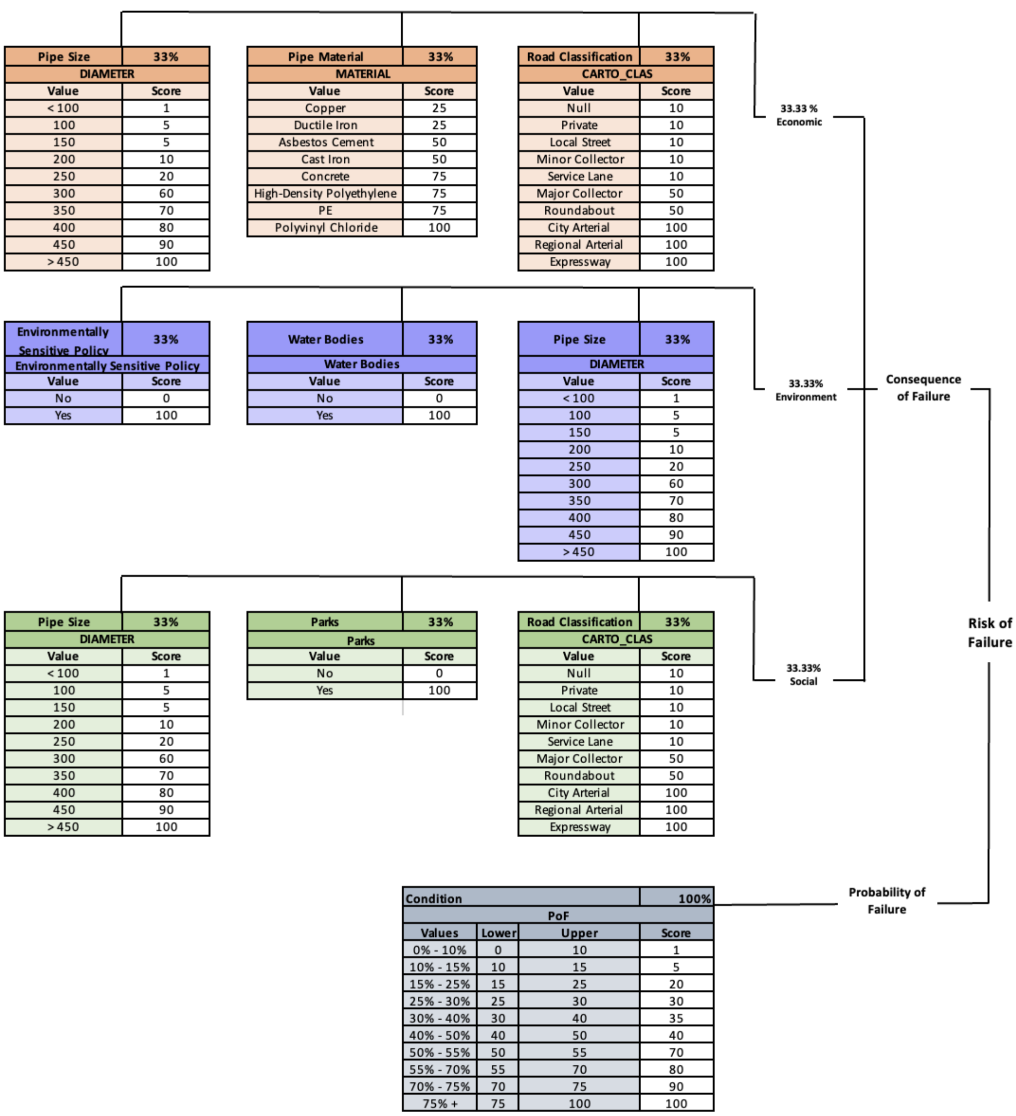
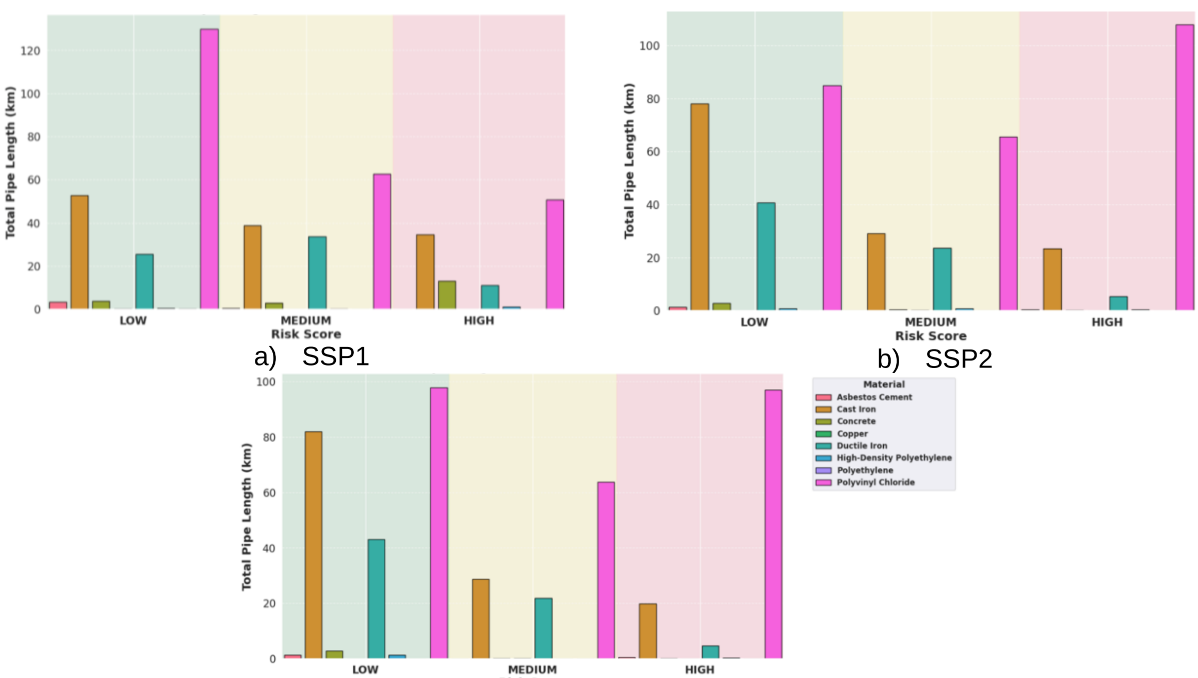


Figure 1: Risk of failure framework

Effective risk assessment of water distribution networks is essential for ensuring infrastructure resilience and long-term sustainability. This section presents the risk assessment results under different climate change scenarios. Water mains are categorized by risk levels across various SSP scenarios. Across all scenarios, a significant portion of water mains fall into the low-risk category, with an average risk score around 5. However, the distribution of low-, medium-, and high-risk pipes varies depending on the climate change scenario. SSP1 has the lowest proportion of high-risk water mains (9%), while SSP2 (15%) and SSP5 (12%) indicate increased vulnerability. The low-risk category is highest in SSP1 (75%), followed by SSP5 (73%), and lowest in SSP2 (69%). SSP1 and SSP2 have the same amount of medium-risk assets (16%), while SSP5 has lower proportion (14%). These results suggest that cooler climate scenarios (SSP1) correspond to lower system vulnerability, whereas warmer scenarios (SSP2 and SSP5) lead to an increase in high-risk water mains. However, the slightly lower proportion of medium- and high-risk pipes in SSP5 compared to SSP2 indicates that some water mains exhibit lower Probability of Failure (PoF) under severe warming conditions. This observation will be explored further in the following section.

Figure 2 presents the material-based length of pipes for different risk scores under three climate change scenarios, highlighting the varying susceptibility of different materials. Cast Iron exhibits the highest high-risk length in SSP1, demonstrating greater vulnerability particularly under cooler climate conditions. This increased risk is likely due to more frequent freeze-thaw cycles, soil movement, and greater susceptibility to corrosion under colder and wetter conditions, aligning with previous findings on climate impacts on water main failures (Khashei, Dziedzic, et al., 2024). PVC, however, is more prominent in the high-risk category in SSP2 and SSP5, indicating that PVC experiences more failures in warmer climate scenarios, particularly under SSP2 and SSP5. This suggests that the vulnerability of Cast Iron is conditionally dependent on climate conditions, with cooler climates exacerbating risk factors, while warmer conditions tend to increase the vulnerability of materials such as PVC. Ductile Iron exhibits a similar pattern to Cast Iron, with more high-risk assets under SSP1. However, under SSP2 and SSP5, the proportion of high- and medium-risk assets decreases, while the low-risk length increases. Another material exhibiting similar behavior to Cast Iron is Concrete, which also has a higher proportion of high-risk assets under SSP1. Since the mentioned materials are more prevalent in this water system, the vulnerability of other materials to different climate change scenarios is not discussed here.



a) SSP1

b) SSP2

c) SSP5

Figure 2: Total pipe length of each material type across risk scores under different climate change scenarios

Furthermore, a spatial analysis of failure risks is conducted under different climate scenarios. To maintain conciseness, spatial distribution of pipe diameters and risk levels for SSP1 is presented in Figure 3. The figure reveals that high-risk pipes are concentrated in areas with relatively larger diameters and along regional and city arterial roads. The failure of these pipes is especially disruptive, as they serve a larger population and are critical to urban transportation, emergency response, and commercial activities. In contrast, pipes located on local streets exhibit lower risk scores, likely due to reduced traffic loads and lower economic consequences in the event of failure. Additionally, pipes are categorized based on road class, and the total pipe length for each category across different risk scores under various climate change scenarios is shown in Figure 4. This categorization helps further understand the combined effects of CoF and PoF on different pipe groups. The three defined categories are as follows: Category 1 includes private roads, local streets, minor collectors, and service lanes; Category 2 includes major collectors and roundabouts; and Category 3 includes city arterials, regional arterials, and expressways.

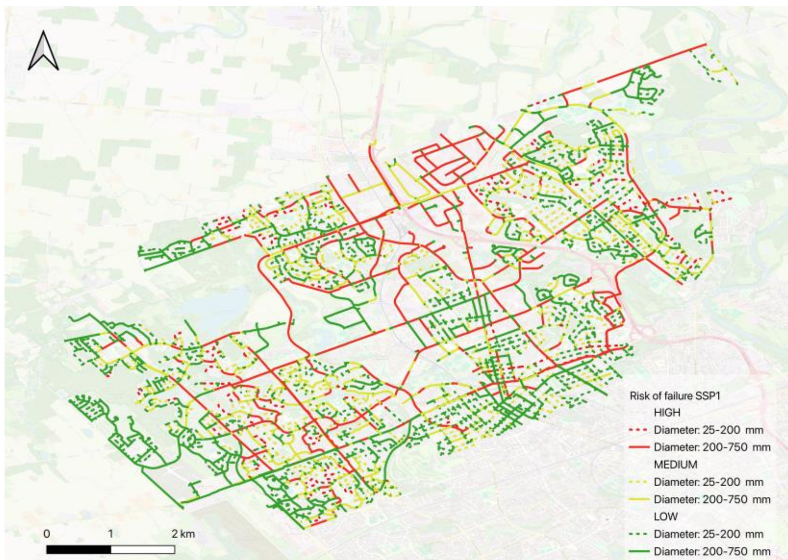
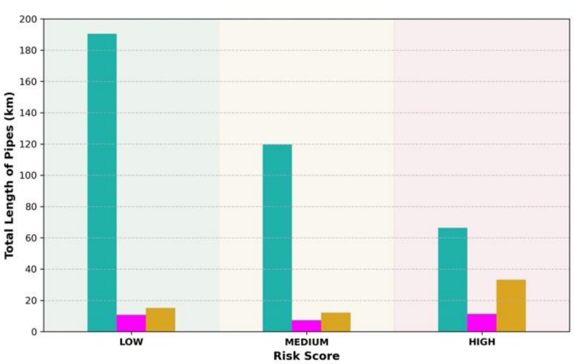
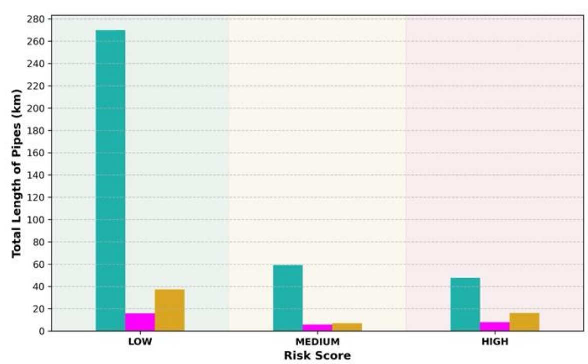


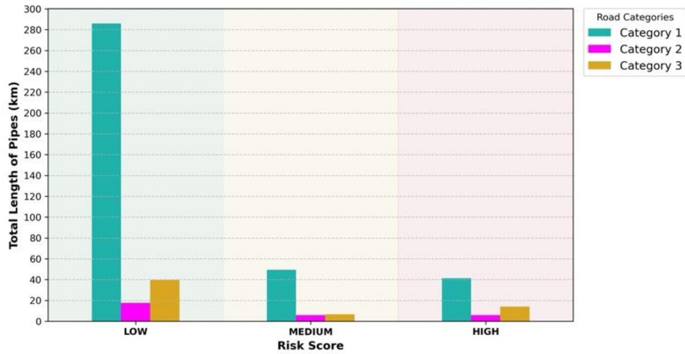
Figure 3: Spatial distribution of risk levels and pipe diameter size in the water mains network- SSP1.



a) SSP1



b) SSP2



c) SSP5

Figure 4: Total pipe length of road categories across risk scores under different climate change scenarios

As seen in Figure 4, although pipes in Category 3 have higher CoF scores, PoF influences the total length of pipes assigned to each risk score under different climate change scenarios. SSP1 results in a higher total length of high-risk pipes, whereas this value decreases under SSP2 and SSP5. This trend is consistent across all road categories. For all categories, the total length of medium- and high-risk pipes decreases from SSP1 to SSP2 and SSP5, while the total length of low-risk pipes increases. These findings emphasize the need for an effective rehabilitation plan to mitigate the impacts of failures, ensuring system resilience and minimizing disruptions to essential urban services.

4. CONCLUSION

This study presents a risk-based prioritization framework for water main replacements, integrating machine learning models and climate change projections to enhance infrastructure resilience. By combining the Probability of Failure and the Consequence of Failure, the framework provides a comprehensive assessment of pipeline vulnerability under different climate scenarios. The case study of the City of Waterloo demonstrates the practical application of this approach, with LightGBM emerging as the most effective predictive model. Findings show that climate variability plays a significant role in pipeline deterioration, with extreme weather events contributing to increased failure risks. Warmer climate scenarios (SSP2 and SSP5) exhibit higher failure probabilities. However, the impact varies across different pipe materials. While PVC pipes show a higher proportion of high-risk assets under SSP2 and SSP5, Cast Iron, Ductile Iron, and Concrete pipes exhibit more high-risk assets under SSP1. The spatial analysis further highlights that high-risk water mains are predominantly located along major pipe diameters, where failures can cause significant disruptions. Incorporating economic, social, and environmental factors into failure consequence assessments ensures a more holistic risk management strategy. This approach allows municipalities to allocate resources efficiently, prioritizing the most critical pipeline segments for rehabilitation.

REFERENCES

Ahmad, T., Shaban, I. A., & Zayed, T. (2023). A review of climatic impacts on water main deterioration. *Urban Climate*, *49*, 101552.

Alshari, H., Saleh, A. Y., & Odabaş, A. (2021). Comparison of gradient boosting decision tree algorithms for CPU performance. *Journal of Institue of Science and Technology*, *37*(1), 157-168.

Almheiri, Z., Meguid, M., & Zayed, T. (2023). Review of Critical Factors Affecting the Failure of Water Pipeline Infrastructure. *Journal of Water Resources Planning and Management*, *149*(10), 03123001.

Barton, N. A., Hallett, S. H., Jude, S. R., & Tran, T. H. (2022). An evolution of statistical pipe failure models for drinking water networks: A targeted review. *Water Supply*, *22*(4), 3784–3813.

Boloukasli ahmadgourabi, F., & Dziedzic, R. (2024). Developing an Open Repository of Water Main Break Prediction Models in Kitchener. *Engineering Proceedings*, *69*(1), 13.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

City of Kitchener. (2024). *Open data portal*. https://open-kitchenergis.opendata.arcgis.com/pages/open-data-listing

City of Waterloo. (2024). *Asset management plan 2024*. https://www.waterloo.ca/en/government/resources/Documents/Finance/Asset-Management/City-of-Waterloo-Asset-Management-Plan.pdf

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21–27.

Creaco, E., Franchini, M., & Alvisi, S. (2012). Evaluating water demand shortfalls in segment analysis. *Water Resources Management*, *26*(8), 2301–2321.

Environment and Climate Change Canada. (2024). *Historical climate data*. https://climate.weather.gc.ca

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958.

Fan, X., Wang, X., Zhang, X., & Yu, X. B. (2022). Machine learning based water pipe failure prediction: The effects of engineering, geology, climate and socio-economic factors. *Reliability Engineering & System Safety*, *219*, 108185.

Fan, X., Zhang, X., Yu, A., Speitel, M., & Yu, X. (2023). Assessment of the impacts of climat change on water supply system pipe failures. *Scientific Reports*, *13*(1), 7349.

Folkman, S. (2018). Water main break rates in the USA and Canada: A comprehensive study. *Mechanical and Aerospace Engineering Faculty Publications*, 174.

Gleick, P. H. (2002). *Dirty-water: estimated deaths from water-related diseases 2000-2020*. Oakland: Pacific Institute for studies in Development, environment, and security.

Godfrey, S., Niwagaba, C., Howard, G., & Tibatemwa, S. (2011). Water Safety Plans for Utilities in Developing Countries-A case study from Kampala, Uganda. *WEDC, Loughborough University. Available online at http://www. lboro. ac. uk/watermark/WEJX7/case-study-annexes. pdf Accessed*, *12*.

Hrudey, S. E., & Hrudey, E. J. (2004). *Safe drinking water*. IWA publishing.

Hussein Farh, H. M., Ben Seghier, M. E. A., Taiwo, R., & Zayed, T. (2023). Analysis and ranking of corrosion causes for water pipelines: a critical review. *NPJ Clean Water*, *6*(1), 65.

Kanta, L., & Brumbelow, K. (2013). Vulnerability, risk, and mitigation assessment of water distribution systems for insufficient fire flows. *Journal of Water Resources Planning and Management*, *139*(6), 593–603.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, *30*.

Khashei, M., Boloukasli ahmadgourabi, F., & Dziedzic, R. (2024). Predicting the Future Failures of Urban Water Systems: Integrating Climate Change and Machine Learning Prediction Models. *Engineering Proceedings*, *69*(1), 35.

Khashei, M., Dziedzic, R., & Roshani, E. (2024). Framework for Predicting Water Main Breaks in the Face of Climate Change. *World Environmental and Water Resources Congress 2024*, 1326–1338.

Latifi, M., Beig Zali, R., Javadi, A. A., & Farmani, R. (2024). Efficacy of Tree-Based Models for Pipe Failure Prediction and Condition Assessment: A Comprehensive Review. *Journal of Water Resources Planning and Management*, *150*(7), 03124001.

MacGillivray, B. H., & Pollard, S. J. T. (2008). What can water utilities do to improve risk management within their business functions? An improved tool and application of process benchmarking. *Environment International*, *34*(8), 1120–1131.

Mamo, T. G. (2015). Risk-based approach to manage aging urban water main infrastructure. *Journal of Water Supply: Research and Technology—AQUA*, *64*(3), 260–269.

Rahman, S., Devera, J., & Reynolds, J. (2014). Risk assessment model for pipe rehabilitation and replacement in a water distribution system. In *Pipelines 2014: From Underground to the Forefront of Innovation and Sustainability* (pp. 1997–2006).

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536.