



## **Integrating machine learning and multi-criteria decision analysis for health risk management in water distribution networks**

Downloaded from: <https://research.chalmers.se>, 2026-05-30 00:36 UTC

Citation for the original published paper (version of record):

Sangroula, U., Vinas, V., Odhiambo, M. et al (2026). Integrating machine learning and multi-criteria decision analysis for health risk management in water distribution networks. *Scientific Reports*, 16. <http://dx.doi.org/10.1038/s41598-026-52465-z>

N.B. When citing this work, cite the original published paper.



## OPEN Integrating machine learning and multi-criteria decision analysis for health risk management in water distribution networks

Uchit Sangroula<sup>1✉</sup>, Victor Viñas<sup>1,2</sup>, Michael Odhiambo<sup>3</sup> & Thomas J. R. Pettersson<sup>1</sup>

Leakages and breaks in water distribution networks (WDNs) cause significant water losses and pose health risks due to pathogen intrusion. The Water Safety Plan (WSP), developed by the World Health Organization (WHO), provides a comprehensive framework for identifying, assessing, and controlling risks within water supply systems. This study demonstrates the application of the WSP framework through a case study of a WDN in Sweden. Pipe break probabilities were estimated using three classification models: Logistic regression, random forest, and extreme gradient boosting (XGBoost), while hydraulic and health consequences were evaluated using hydraulic modelling and Quantitative Microbial Risk Assessment (QMRA) to quantify the overall health risk. A Multi-Criteria Decision Analysis (MCDA) approach, specifically the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) was utilized to prioritize risk mitigation strategies through breakage and leakage control measures. The proposed approach integrates predictive modelling, consequence evaluation, and decision analysis, offering a structured method for water utilities in prioritizing interventions and improving the overall safety and reliability of WDNs.

**Keywords** Water distribution networks, Water safety plan, Health risks, Decision making

Water distribution networks (WDNs) are vital components of water supply systems necessary for providing consumers with sufficient water of appropriate quality<sup>1</sup>. Physical defects such as pipe breaks, cracks, and leakages are characteristic of aging water distribution infrastructure. Leakages and breaks in distribution pipelines result in significant water losses, causing substantial financial, environmental, and social impacts<sup>2</sup>. Globally, the volume of water lost due to leakages and breakages is estimated at 346 million cubic meters per day<sup>3</sup>. In Sweden, approximately 15–20% of the municipal drinking water is lost due to leakage and breakages<sup>4</sup>. Additionally, pipe breaks and leakages can serve as pathways for pathogen intrusion due to inadequate pressure, posing serious risks to human health<sup>5</sup>. Practical rehabilitation strategies are therefore necessary to maintain distribution pipelines and reduce associated losses and health risks<sup>6</sup>. Developing such strategies requires considering risks related to pipe breaks and leakage, budget constraints, and resource limitations.

The World Health Organization (WHO) has recognized the importance of ensuring water safety in distribution systems and subsequently developed the Water Safety Plan (WSP) framework<sup>7</sup>. The WSP encompasses a comprehensive risk assessment and management approach, where potential hazards within the water supply system are identified, and control measures implemented to prevent contamination. The principal objective of WSPs is to proactively protect public health by maintaining water quality from source to consumers. The preventive strategy of WSPs, which focuses on identifying and mitigating risks, has been recommended by WHO guidelines for the preventive management of WDNs<sup>8</sup>. WSPs have been successfully implemented worldwide, demonstrating significant benefits in reducing health risks<sup>9</sup>.

There is also an increasing need for integrated, data-driven, and stakeholder-inclusive approaches that enable water utilities to identify significant risks in their WDNs and select the most effective interventions. Previous studies have typically utilized multiple tools independently for prediction, risk assessment, and maintenance planning in WDNs. For example, some studies have employed machine learning (ML) for pipe break predictions<sup>10</sup>, others have applied Quantitative Microbial Risk Assessments (QMRA) for contamination evaluations<sup>11</sup>, and

<sup>1</sup>Department of Architecture and Civil Engineering, Water Environment Technology, Chalmers University of Technology, Gothenburg SE-412 96, Sweden. <sup>2</sup>AFRY AB, Grafiska vägen 2A, Gothenburg SE-412 83, Sweden. <sup>3</sup>Department of Urban Water Engineering, Norconsult AB, Gothenburg SE-402 76, Sweden. ✉email: uchit@chalmers.se

some have utilized decision-making methods for hazard prioritization<sup>12</sup>. However, an integrated approach that combines these elements remains limited in the literature.

This study proposes a framework that integrates predictive modelling, consequence evaluation, and decision analysis within the WSP framework. The study involves four distinct steps: (i) identifying vulnerable pipes and assessing their impact upon failure, (ii) conducting health risk assessments, (iii) identifying control measures, and (iv) using Multi-Criteria Decision Analysis (MCDA) to prioritize leakage and breakage control strategies for health risk management. The likelihood of pipe breaks (P) is estimated using multiple supervised ML classification models<sup>13,14</sup>. Consequences (C) are assessed by evaluating impacts on both human health and network performance, by combining QMRA<sup>15</sup> and hydraulic modelling (EPANET 2.2)<sup>16</sup>. Risk assessment is performed by multiplying the probability (P) by the consequence (C). A GIS-based risk map is then developed to categorize pipes according to their risk levels within the network.

Subsequently, a series of interviews are conducted with water experts and stakeholders for the identification of different strategies for health risk reduction. A MCDA approach<sup>17</sup> is utilized to determine the best strategies with regards to reducing health risks. The Technique for Order Preferences by Similarity to Ideal Solutions (TOPSIS) method<sup>18</sup> is utilized for the MCDA. This approach is demonstrated through a case study on a local WDN in Skåne, Sweden. The novelty of this work lies in uniting predictive modelling, microbial risk analysis, and multi-criteria decision support into one integrated framework. The results from this study can help water utilities prioritize pipe renewal and rehabilitation and proactively address vulnerabilities and health risks in WDNs.

## Background

### *Risk assessment*

A risk assessment is a structured process that identifies both the probability of occurrence of an undesired event and the extent of the adverse consequences that may arise from that event<sup>19</sup>. For a WDN risk assessment, the general term for risk is expressed as Eq. (1).

$$R = P \cdot C \quad (1)$$

where R is the risk, P is the likelihood of occurrence of an undesired event (pipe break) and C represents the consequences of the event (associated impacts).

**Probability of pipe breaks** Numerous studies have investigated pipe break probabilities in WDNs<sup>20,21</sup>, categorizing pipe break causes into physical, environmental, and operational factors<sup>22</sup>. Common physical factors include pipe length, diameter, material, and age. Environmental factors often include traffic loads, soil types, and climate conditions, whereas operational factors include internal pressure, past failures, and pump operations. Variables predicting pipe breaks are termed explanatory variables.

In general, three main prediction model types are utilized: physical<sup>23</sup>, statistical<sup>24</sup>, and ML models<sup>25</sup>. Physical models simulate mechanical and environmental stresses on pipes, but require extensive and costly pipe inspection data (field measurements)<sup>26</sup>. Statistical and ML models, being more cost-effective and data-efficient, predict failures based on historical data and mathematical relationships<sup>27</sup>. Recently, ML models have grown in popularity due to superior predictive accuracy and robust handling of complex data relationships<sup>28,29</sup>. Typical model outputs include failure probabilities, pipe classifications, time to failure, and failure rates.

**Consequences of pipe breaks** Pipe breaks can result in a broad range of serious consequences such as water loss, high repair, and replacement costs, threats to human health, environmental damage and customer dissatisfaction<sup>30</sup>. Hydraulic and health consequences are particularly important as they directly affect system reliability and public safety. Hydraulic consequences involve network disruptions affecting water delivery, measured by disconnected nodes and unsupplied flow volume. Health consequences involve contamination risks, potentially leading to waterborne diseases<sup>31–33</sup> and endemic illnesses globally<sup>34,35</sup>. Contaminants such as *Campylobacter*, *Salmonella*, and norovirus from pipe breaks can cause illnesses ranging from mild to severe<sup>36</sup>.

### *Multi-criteria decision analysis (MCDA)*

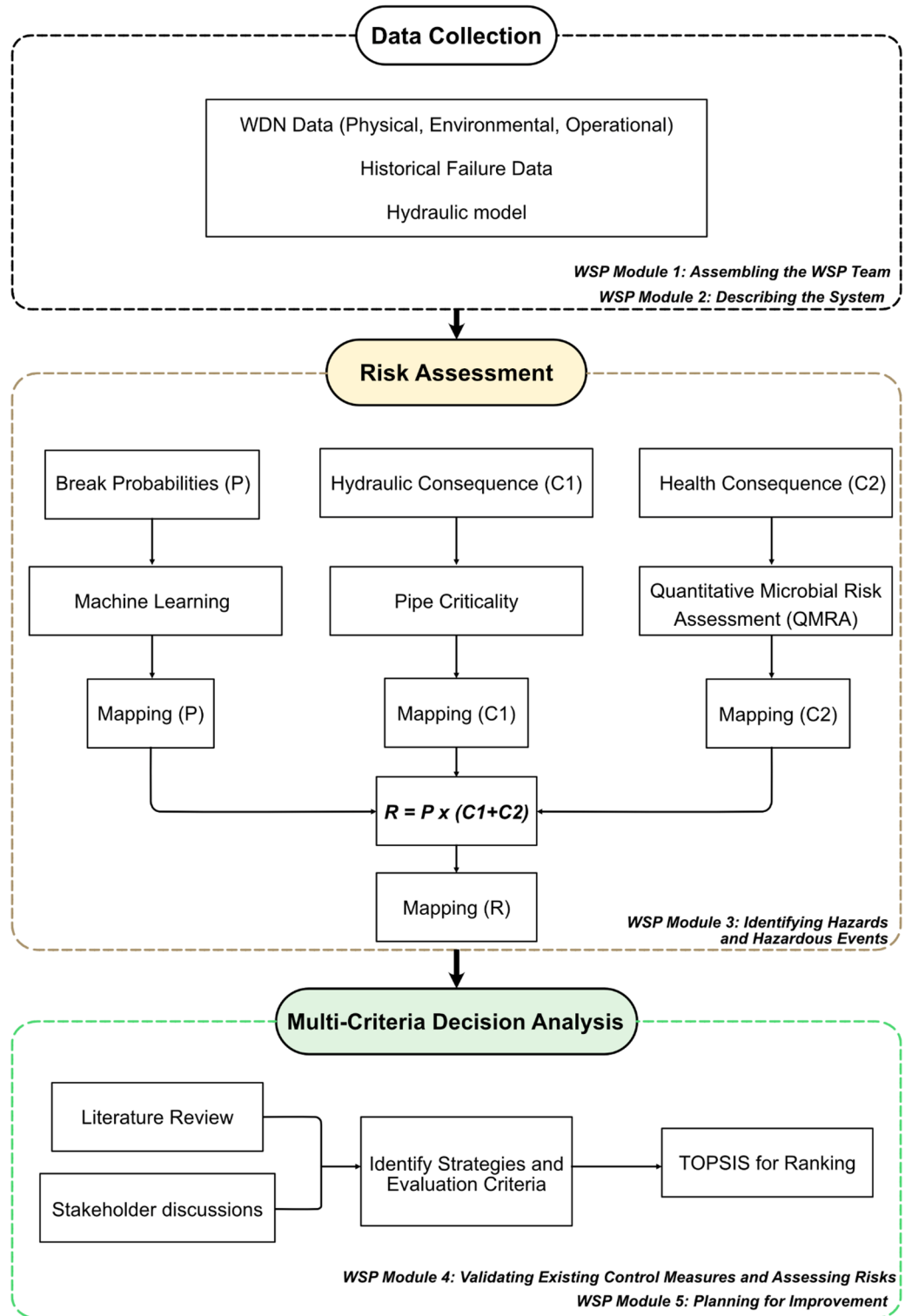
Decision making is an integral part of any successful project. It can involve either individuals or groups and can be based upon a single criterion or take multiple criteria into account<sup>37</sup>. MCDA is a structured decision-making process which aims to prioritize alternatives while considering multiple varying criteria<sup>38</sup>. The typical MCDA process involves: (i) defining the decision problem, objectives, alternatives, and evaluation criteria; (ii) structuring the problem; (iii) incorporating stakeholder preferences; (iv) evaluating alternatives; and (v) analysing results.

MCDA methods can be categorized into Multi-Objective Decision Making (MODM), focusing on optimizing continuous variables, and Multi-Attribute Decision Making (MADM), which selects optimal discrete alternatives. Both methods have found extensive utility in the field of water supply management, including the design of distribution systems<sup>39</sup>, optimization of monitoring networks<sup>40</sup>, and maintenance planning to minimize risks<sup>41,42</sup>. Recent studies have highlighted the superior effectiveness of MADM methods over MODM approaches in water distribution system rehabilitation<sup>43,44</sup>. Common examples of MADM methods utilized

for decision-making tasks include the Analytical Hierarchy Process (AHP)<sup>45</sup>, TOPSIS<sup>18</sup>, and Elimination and Choice Translating Reality (ELECTRE)<sup>46</sup>.

**Methods**

The WSP framework consists of 11 key modules where each module builds upon the previous step. Figure 1 illustrates the specific WSP modules emphasized in this study.



**Fig. 1.** Overall structure of the study adapted from the WSP framework<sup>7</sup>.

### Case study description

This study focused on a small segment of a larger WDN located in Skåne County, southern Sweden, previously analysed by Odhiambo, et al. [11], hereafter referred to as the case study section.

This section spans 13.5 km, with pipe diameters ranging from 25 mm to 200 mm. The network is dominated by polyethylene (PE), which accounts for 41% of the total pipe length, followed by cast iron (36%) and ductile iron (11%), while the remaining materials together account for 12%. Based on the historical break record, the overall break rate in the case study section was 13.4 breaks/100 km/year. Among the dominant materials, ductile iron exhibited the highest break rate (24.3 breaks/100 km/year), followed by cast iron (18.5 breaks/100 km/year) and PE (10.2 breaks/100 km/year).

The annual water supply volume is approximately 50,000 m<sup>3</sup>, with losses due to pipe breaks and leakages estimated at 0.054 l/s/km, representing 27% of total water production. The pipes in this WDN are installed at depths of 1.5 to 2 m below the ground surface. Local staff indicated that most sewer pipes are positioned at the same elevation as drinking water pipes, with a horizontal separation of approximately 0.5 m.

### Risk assessment methodology

The risk assessment was done by considering both the likelihood of pipe breaks and their hydraulic and health consequences, as outlined in Eq. (2). Detailed descriptions of each step are provided in the sections below.

$$Risk = Likelihood\ of\ Break \cdot (Hydraulic\ consequence + Health\ consequence) \quad (2)$$

#### Break probabilities

Three ML models: logistic regression, random forest, and eXtreme Gradient Boosting (XGBoost) were used to estimate the probability of pipe break. Brief descriptions of these models are provided below.

**Logistic regression** Logistic regression (LR) estimates the probability of a binary outcome by applying a logistic (sigmoid) function to a weighted sum of the input features. LR is mathematically formulated as Eq. (3) [47].

$$p = \frac{1}{1 + e^{-(w_0 + \sum_{i=1}^m w_i \cdot x_i)}} \quad (3)$$

where  $p$  is the output probability of each sample,  $x_i$  denotes the value of the  $i^{th}$  feature,  $w_i$  is the weight of the  $i^{th}$  feature and  $w_0$  is the bias constant. The weighted sum of the features, combined with the bias, is passed through the sigmoid function, which transforms it into a value between 0 and 1; this value is then compared to a decision threshold to assign each pipe to the intact or break class.

**Random forest** Random forest (RF) is an ensemble learning algorithm that estimates the probability of a binary outcome by constructing many decision trees and aggregating their predictions [48]. It is a bagging-based method, where each tree is trained on a different bootstrap sample of the training data, and at each split a random subset of features is considered. The predicted probability for an input feature vector  $x$  is obtained by averaging the probabilities from  $B$  trees, as shown in Eq. (4):

$$p = \frac{1}{B} \sum_{b=1}^B p_b(x) \quad (4)$$

where  $p$  is the output probability for each pipe,  $p_b(x)$  denotes the probability predicted by the  $b^{th}$  decision tree for the input feature vector  $x$ , and  $B$  is the total number of trees in the forest. The probabilities from all trees are averaged to produce a final value between 0 and 1; this value is then compared to a decision threshold to assign each pipe to the intact or break class.

**XGBoost** Extreme gradient boosting (XGBoost) is an ensemble learning algorithm that estimates the probability of a binary outcome by building decision trees sequentially, where each new tree is trained to improve the predictions of the previous trees<sup>49</sup>. The model combines the outputs of  $T$  trees into a single score and then converts this score into a probability using the logistic (sigmoid) function, as shown in Eq. (5):

$$p = \frac{1}{1 + e^{-f(x)}}, \quad f(x) = \sum_{t=1}^T g_t(x) \quad (5)$$

where  $p$  is the output probability for each pipe,  $x$  is the input feature vector,  $g_t(x)$  is the contribution of the  $t^{th}$  decision tree, and  $f(x)$  is the combined score from all trees. The sigmoid function transforms the score into a value between 0 and 1; this value is then compared to a decision threshold to assign each pipe to the intact or break class.

**Data description and preparation** The explanatory variables (i.e., model features) are selected from three different categories i.e. physical, environmental, and operational. Table 1 provides an overview of the explanatory variables adopted for the model.

Physical pipe information and break records provided by the municipality were spatially matched using geoprocessing. The average operational pressure was simulated over a 24-hour period. Soil data were obtained from the Geological Survey of Sweden (SGU), the national authority for geology and related geosciences.

Explanatory variable	Name	Type	Description
Physical	Length	Numerical	Pipe length in m
	Diameter	Numerical	Pipe diameter in mm
	Age	Numerical	Pipe age in years
	Material	Categorical	Material of pipe section
Environmental	Elevation	Numerical	Elevation from mean sea level
	Soil type	Categorical	Major soil type at pipe location
	Roads & Trainlines	Categorical	Proximity to the pipe
Operational	Pressure	Numerical	Average operational pressure
	Impact score	Numerical	Degree to which a pipe is affected by failures of other pipes in WDN
	Historical breaks	Numerical	Recorded breaks in the last 20 years

**Table 1.** Explanatory variables for the ML models.

Proximity to infrastructure (roads and train lines) was determined by creating buffer zones around pipes and overlaying these with relevant spatial datasets.

In this study, an ‘Impact score’ feature was developed to capture how significantly each individual pipe is affected by the failures occurring elsewhere in the distribution network. To achieve this, a comprehensive hydraulic modelling analysis was conducted, in which failure was simulated in each pipe of the network systematically. For each simulated pipe failure, the resulting changes in pressure and flow across the entire distribution network were monitored. This analysis enabled the quantification of how sensitive each pipe is to hydraulic disturbances caused by failures occurring elsewhere in the network. The cumulative hydraulic impact experienced by each pipe across all simulations was normalized and squared to derive the ‘Impact score.’ Pipes receiving higher impact scores indicate those pipes which are more frequently and severely affected by failures occurring elsewhere in the network.

Previous studies that have addressed spatial or network-level dependencies in a WDN by employing clustering methods or hotspot analyses, such as spatial clustering or k-means segmentation<sup>50,51</sup>, have relied on historical failure patterns or physical characteristics. In contrast, the impact score dynamically captures the spatial relationships and interdependencies among pipes within the WDN.

The ‘Historical breaks’ variable was used to construct the ‘target’ variable for the model, enabling an assessment of its predictive performance. The model was designed to predict whether a given pipe experienced at least one failure during the study period of 20 years. The dataset obtained from the utility included the entire distribution network in Skåne County, Sweden. However, the training and testing datasets were carefully separated based on spatial boundaries. The ML models were trained on the network data, explicitly excluding the case study section, which was specifically set aside for testing. Hence, the model was tested only on the case study section of the distribution network. A notable idea explored by Daulat, et al.<sup>52</sup>, investigated whether ML models trained on data from large or multiple utilities could be used to predict pipe failures in smaller, separate utilities. The approach used in this study adopts a similar idea within a single WDN by spatially separating the training and testing data.

The dataset was preprocessed by removing missing/erroneous records, where approximately 6.5% of records were removed, mainly due to missing core inventory attributes (e.g., installation year, diameter, or length). Inconsistent observations (e.g., failures recorded before installation) and extreme outliers (e.g., diameter = 0) were also filtered. Moreover, since pipe inventory and break records were spatially matched, a small number of break events that could not be matched to any pipe segment were excluded.

Numerical variables were normalized using z-score standardization, with scaling parameters estimated from the training data only to avoid information leakage. A total of three categorical variables (material, roads and trainlines, and soil type) were one-hot encoded<sup>53</sup>. WDNs commonly exhibit imbalanced datasets, as majority of pipes have never experienced a break. This imbalance makes model training difficult because the model struggles to learn patterns associated with the minority class (pipes that have experienced breaks) resulting in poor predictive performance<sup>13,54</sup>. In the present dataset, the positive (break) class represented approximately 7% of samples. To balance the dataset, random oversampling method<sup>55</sup> was applied only on the training dataset.

**Model training, hyperparameters tuning and selection** Model training and hyperparameter tuning were performed only on the training dataset using stratified *k*-fold cross-validation to maintain a similar break/non-break class proportion in each fold. Hyperparameters for LR, RF, and XGBoost were optimized using randomized search with area under the precision-recall curve (PR-AUC) as the primary scoring metric, which is more suitable than accuracy for imbalanced classification. Model selection was performed using training-only out-of-fold (OOF) PR-AUC to avoid test-set bias. The selected model was then refit on the full training set and evaluated once on the test case study section. The hyperparameter search ranges are summarized in Table 2, and the complete search ranges are reported in Supplementary Table S1.

**Metrics for model evaluation** Model performance was evaluated using both threshold-independent and threshold-dependent metrics. Threshold-independent metrics were used for model benchmarking and selection and included the area under the receiver operating characteristic curve (ROC-AUC) and the PR-AUC. ROC-AUC summarizes class discrimination across all probability thresholds, while PR-AUC summarizes the

Model	Hyperparameters	Type	Range
LR	Solver	Categorical	lbfgs, saga
	Penalty	Categorical	lbfgs: l2; saga: l1, l2, elasticnet
	Regularization strength	Float (log-uniform)	$[10^{-4}, 10^2]$
	L1 Ratio (if elasticnet)	Float (uniform)	[0.05, 0.95]
RF	Number of estimators	Integer	[30, 300]
	Maximum depth	Integer	[2, 6]
	Maximum features	Categorical	sqrt, log2
	Minimum samples leaf	Integer	[1, 8]
XGBoost	Number of estimators	Integer	[30, 300]
	Learning rate	Float (log-uniform)	[0.01, 0.1]
	Maximum depth	Integer	[2, 6]
	Feature subsampling	Float (uniform)	[0.7, 1.0]

**Table 2.** Summary of hyperparameters of ML models.

		Predicted Status	
		Predicted status is 'Break'	Predicted status is 'Intact'
Actual Status	Actual status is 'Break'	True Break (TP)	False Intact (FN)
	Actual status is 'Intact'	False Break (FP)	True Intact (TN)

**Fig. 2.** Confusion matrix for pipe status classification.

precision-recall trade-off and is particularly informative under class imbalance. To quantify uncertainty, 95% confidence intervals for ROC-AUC and PR-AUC were estimated on the case study test set using bootstrap resampling.

Threshold-dependent metrics were computed from the confusion matrix (Fig. 2), which can be used to evaluate the model's performance<sup>56</sup>. A confusion matrix contains the number of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) samples as a result of comparing the true and the predicted labels for all the test samples. The term positive is associated with class 1 'break' and the term negative with class 0 'intact'.

Based on the confusion matrix, the performance of the model can be evaluated using metrics such as accuracy, precision, recall, and F1-score. However, for datasets with class imbalance, accuracy may not provide a reliable measure of model performance, as it can be heavily biased toward the majority class. Hence, precision, recall, and F1-score are more reliable metrics for evaluating the model performance. The metrics are defined below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = Sensitivity = \frac{TP}{TP + FN} \quad (8)$$

$$F1 \text{ score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (9)$$

#### Hydraulic consequence

The Water Network Tool for Resilience (WNTR) version 1.3.2<sup>57</sup> was used to simulate and analyse the hydraulic impacts of pipe breaks within the WDN. Hydraulic consequence analysis was performed using the WNTR

Simulator under pressure-driven demand (PDD) conditions over a 24-hour period and the Wagner equation<sup>58</sup> was used to determine nodal demands based on available pressure.

$$qD_i = \begin{cases} D_i, & p_i \geq P_f, \\ D_i \left( \frac{p_i - P_0}{P_f - P_0} \right)^{\frac{1}{\varepsilon}}, & P_0 < p_i < P_f, \\ 0, & p_i \leq P_0 \end{cases} \quad (10)$$

where  $qD_i$  is the actual demand at node  $i$ ,  $D_i$  is the nominal demand at node  $i$ ,  $p_i$  is the pressure at node  $i$ ,  $P_0$  is the minimum pressure below which the demand is zero,  $P_f$  is the pressure required to fully satisfy the demand  $D_i$ , and  $1/\varepsilon$  is the pressure function exponent usually set to 0.5 modelling flow through orifices or pipes<sup>16</sup>.

Hydraulic consequence was measured by enumerating nodes experiencing pressure below a critical threshold (1 m) after each pipe closure simulation. Nodes already experiencing low baseline pressure were excluded. Results from these simulations were normalized and visualized, highlighting critical pipes whose failures cause widespread low-pressure scenarios.

#### QMRA and health consequence

The QMRA and health consequence analysis were also conducted using WNTR version 1.3.2. The network model was configured with a duration of 48 h, a hydraulic timestep of 1 h, and a quality timestep of 5 min. The demand model was set to PDD. For each pipe in the network, a break was simulated individually by splitting the pipe and introducing a local intrusion point at the new node. No pipe closures or segment-level isolations were applied; thus, only one pipe was affected at a time while adjacent pipes remained in service. The intrusion volume was estimated using the orifice-based formulation in Eq. (11) based on nodal pressure and orifice discharge parameters.

$$d_{leak} = C_d A p^\alpha \sqrt{\frac{2}{\rho}} \quad (11)$$

where  $d_{leak}$  is the leak demand,  $C_d$  is the discharge coefficient,  $A$  is the area of the hole,  $p$  is the gauge pressure inside the pipe,  $\alpha$  is the pressure exponent, and  $\rho$  is the density of the fluid. The discharge coefficient was set to 0.6. The pressure exponent was set at the common value of 0.5.

Random values for dilution factor ( $D$ ), frequency factor ( $F$ ), and pathogen concentration in faeces ( $C_p$ ) were generated using specified distributions: uniform ( $10^{-2}$ ,  $10^{-4}$ ) for dilution factor, uniform (0.01, 1) for frequency and triangular ( $10$ ,  $10^2$ ,  $10^3$ ) for pathogen concentration in faeces. The intrusion mass ( $M$ ) was then calculated as:

$$M = V \cdot C_p \cdot D \cdot F \quad (12)$$

The mass was introduced into the downstream node of the broken pipe, and a hydraulic and water quality simulation was performed. A conservative tracer approach was adopted, meaning that the contaminant was transported without decay or reaction with disinfectant residuals, representing a worst-case contamination scenario. The results were analysed to determine contamination levels at various nodes.

After the quality simulation, the probability of infection ( $P_{inf}$ ) is calculated based on the contamination levels. The dose ( $D$ ) is calculated as the product of concentration and volume

$$D = C_{node} \cdot V \quad (13)$$

where  $D$  is the dose,  $C_{node}$  is the pathogen concentrations at the consumer tap and  $V$  is the volume of water ingested per person.

The Exact Beta-Poisson model was used to assess the dose-response for each reference pathogen. Using this dose-response model, the probability of infection was calculated with:

$$P_{inf} = 1 - \exp^{-r \cdot D} \quad (14)$$

where  $P_{inf}$  is the probability of infection,  $r$  is a sample from the Beta distribution with  $\alpha$ ,  $\beta$  parameters for each pathogen and  $D$  is the dose ingested.

Pathogen intrusion concentrations were based on E. coli levels as indicators of faecal pollution, using established literature and field data<sup>59,60</sup>. Reference pathogens chosen were *Campylobacter*, norovirus, and *Cryptosporidium*, which represent typical Swedish WDN risks<sup>61</sup>.

The Beta distribution parameters for each reference pathogen were: *Campylobacter* ( $\alpha = 0.024$ ,  $\beta = 0.011$ )<sup>62</sup>, norovirus ( $\alpha = 0.04$ ,  $\beta = 0.055$ )<sup>63</sup> and *Cryptosporidium* ( $\alpha = 0.115$ ,  $\beta = 0.176$ )<sup>64</sup>. The population was assumed to be homogenous, i.e., all consumers would respond according to the chosen dose-response function. Results were aggregated into mean risks and percentiles, enabling ranking of pipes by infection risk severity.

#### TOPSIS for decision making

**Development of decision strategies and evaluation criteria** Following the risk assessment, several strategies for reducing health risks through breakage and leakage control were identified based on extensive literature review and structured stakeholder discussions. Ten industry experts, comprising water specialists, utility managers, and water supply consultants, participated in the evaluation. These experts were selected based on their experience and expertise in water distribution system maintenance and rehabilitation in Sweden. The study was carried out

Strategies	Strategy description
Pressure management	Regulate water pressure through pressure-reducing valves.
Pipe repair and rehabilitation	Conduct repairs and apply techniques such as pipe lining.
Pipe replacement/renewal	Replace old or deteriorating pipes with new ones (pipe renewal).
Increase inspection and testing	Enhance the frequency and thoroughness of pipe inspections.
Increase metering	Install additional meters to better monitor water flow and usage.
Active leakage control	Implement strategies for continuous measurement and management of leaks, including the use of specialized tools.
Cross connection control	Implement measures to identify and manage unauthorized or improper cross-connections.
Public awareness and engagement	Educate and engage the community on the importance of leak reporting.

**Table 3.** Different strategies for reducing risks in WDNs.

Evaluation criteria (E)	Criteria description
Cost	The financial expenditure required to implement and maintain the strategy.
Executability	The ease and feasibility of implementing the strategy within the existing infrastructure.
Risk reduction	The extent to which the strategy lowers the overall risk.
Social concern	The strategy's impact on the community and stakeholders (social, environmental and public health).
Reliability	The consistency and dependability of the strategy over time.

**Table 4.** Evaluation criteria for different strategies.

in accordance with relevant guidelines and regulations. Ethical approval was not required under the Swedish ethical review framework, as the stakeholder input involved only professional judgments and did not involve any intervention or the collection/processing of sensitive personal data. Informed consent was obtained from all participants.

To facilitate the evaluation process, each expert received an evaluation spreadsheet detailing the various risk-reduction strategies alongside clearly defined evaluation criteria. The identified strategies are shown in Table 3 and the evaluation criteria are shown in Table 4. Stakeholders were requested to provide their judgment by assigning scores on a scale ranging from 1 to 5. On this scale, a score of 1 indicated that the expert considered the strategy the least favourable with respect to the specific criterion, whereas a score of 5 indicated the highest favourability. The expert evaluations were used to prepare a decision matrix, where each row represents a strategy, and each column represents an evaluation criterion.

$$Decision\ Matrix = \begin{pmatrix} Strategy_1 & E_1 & E_2 & E_3 & E_4 & E_5 \\ Strategy_1 & E_1 & E_2 & E_3 & E_4 & E_5 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Strategy_8 & E_1 & E_2 & E_3 & E_4 & E_5 \end{pmatrix}$$

**TOPSIS analysis** The TOPSIS method was applied to rank the identified strategies. TOPSIS identifies the best alternative as the closest to an Ideal Solution (IS) and furthest from a Negative Ideal Solution (NIS). The analysis involved the following steps:

**Normalization of the decision matrix** Criteria scores were normalized using vector normalization to ensure comparability:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}} \quad (15)$$

where  $x_{ij}$  is the score of strategy  $i$  for criterion  $j$ ,  $n$  is the number of strategies and  $r_{ij}$  is the normalized value.

**Calculation of ideal and negative ideal solutions** The IS and NIS are defined as follows:

$$v^+ = (\max(v_{1j}), \max(v_{2j}), \dots, \max(v_{nj})) \quad (16)$$

$$v^- = (\min(v_{1j}), \min(v_{2j}), \dots, \min(v_{nj})) \quad (17)$$

In Eq. (16) and Eq. (17),  $v_{ij}$  is the value for strategy  $i$  and criterion  $j$ ,  $v^+$  is the ideal solution, taking the maximum value for each criterion across all strategies and  $v^-$  is the negative ideal solution, taking the minimum value for each criterion across all strategies.

**Calculation of Euclidean distances** For each strategy, the distance from both the ideal best and ideal worst solutions was calculated using Euclidean distance.

$$S_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2} \quad (18)$$

$$S_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2} \quad (19)$$

where  $S_i^+$  and  $S_i^-$  are the distances from the IS and NIS, respectively.

**TOPSIS score calculation** The TOPSIS score for each strategy is calculated using the Relative Closeness ( $RC_i$ ), which measures how close a strategy is to the IS while accounting for its distance from the NIS. Strategies with higher TOPSIS scores represent more favourable alternatives.

$$RC_i = \frac{S_i^-}{S_i^+ + S_i^-} \quad (20)$$

To assess uncertainty in the MCDA inputs, a bootstrap analysis was performed by resampling the experts with replacement and recomputing the normalized criterion weights and corresponding TOPSIS scores for each replicate. The resulting bootstrap distributions were used to report 95% uncertainty intervals for the criterion weights and TOPSIS scores.

## Results and discussions

### Probability assessment of the case study

The first step of the analysis involved identifying vulnerable pipes by utilizing the ML models. As described in Sect. "Break Probabilities", all ML models were trained while explicitly excluding the case study section, which was reserved for evaluation purposes. Based on the training-only OOF PR-AUC criterion, RF was selected as the final predictor. Table 5 summarizes the discrimination performance of all ML models on the case study section using ROC-AUC and PR-AUC, with 95% bootstrap confidence intervals.

Due to security reasons, detailed visualizations and performance results are presented only for the case study section, summarized in Table 6, which reports the corresponding confusion matrix and threshold-dependent performance metrics for the RF model. The decision threshold for RF (0.36) was chosen using training-only OOF predictions to meet a minimum precision of 0.50 and was then applied unchanged to the case study section.

Despite the limited test set size, the RF model showed reasonable performance in classifying failure-prone pipes. The model was able to identify 60% of the actual pipes that suffered a break during the study period over the past 20 years. Similarly, the model was able to classify 96% of the pipes not experiencing a break. The model precision was 54%, accuracy 93%, and the resulting F1-score was 57%.

To provide additional insight, permutation importance was evaluated for the RF model as the decrease in average precision resulting from randomly permuting each input feature. The results showed that age was the most influential variable, followed by material and length. This is consistent with findings reported in previous studies, which have similarly identified physical pipe characteristics as the most influential features in failure prediction<sup>65-67</sup>. Soil type, Impact score, diameter, and traffic intersection exhibited lower and broadly similar importance. These findings indicate that the model was driven mainly by physical pipe-related characteristics, although the observed importance patterns should be interpreted in the context of the relatively small case study section.

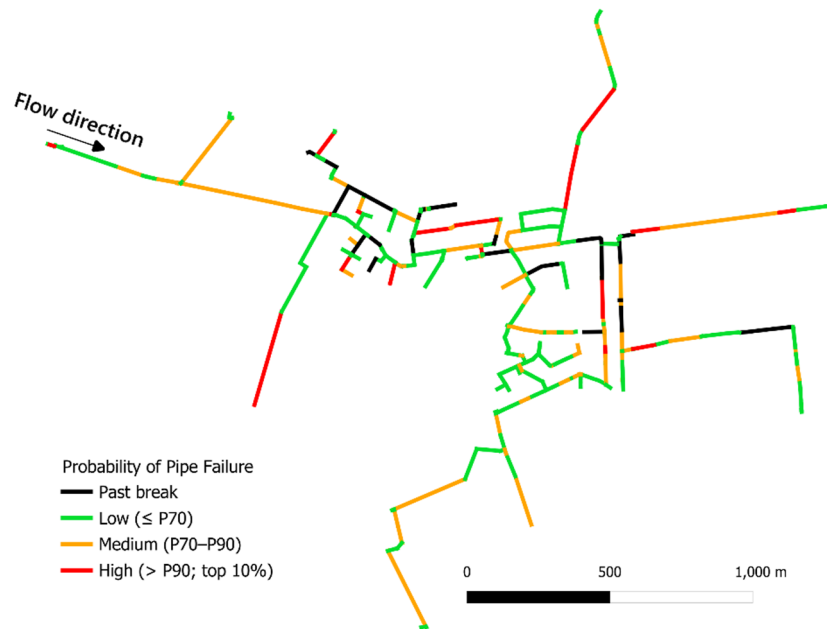
The model's output represents the model's confidence in classifying a pipe as belonging to the category of historically failed pipes. For pipes with no prior failures, a high predicted probability indicates that they share similar characteristics with failed pipes, suggesting increased vulnerability and a greater risk of future breaks. While these probability values do not represent absolute failure predictions, they provide a quantitative basis for

Model	ROC-AUC (95% CI)	PR-AUC (95% CI)
RF	0.897 [0.82, 0.95]	0.645 [0.48, 0.79]
XGB	0.883 [0.79, 0.95]	0.635 [0.47, 0.78]
LR	0.858 [0.75, 0.94]	0.595 [0.43, 0.73]

**Table 5.** Performance of ML models on the case study section.

		Predicted status	
		Predicted 'break'	Predicted 'Intact'
Actual Status	Actual 'Break'	21	14
	Actual 'Intact'	18	402
Evaluation metrics of the RF model (with threshold: 0.36)			
Accuracy	Precision	Recall	F1-score
0.93	0.54	0.6	0.57

**Table 6.** Confusion matrix and evaluation metrics on the case study section.



**Fig. 3.** Map of the WDN showing RF-predicted break vulnerability. Pipes are grouped into relative vulnerability tiers based on the distribution of predicted probabilities: lower vulnerability ( $\leq P70$ ; bottom 70%), moderate vulnerability (P70–P90; next 20%), and higher vulnerability ( $> P90$ ; top 10%). Past breaks are shown in black.

assessing relative vulnerability across the network, thereby supporting the prioritization of maintenance efforts. Improving the model's predictive accuracy was not the primary focus of this study, as the model was intended to serve just as an intermediate step within the broader risk assessment framework.

The predicted break probabilities were calibrated so that the numerical probability values better reflect observed event frequencies. Specifically, the RF output scores were post-processed using a sigmoid calibration model fitted on the training data only (via cross-validation) and then applied unchanged to the case study section. The calibrated probabilities were subsequently combined with consequence values to calculate overall pipe-level risk. Furthermore, the results were visualized using a GIS-based map to support spatial interpretation.

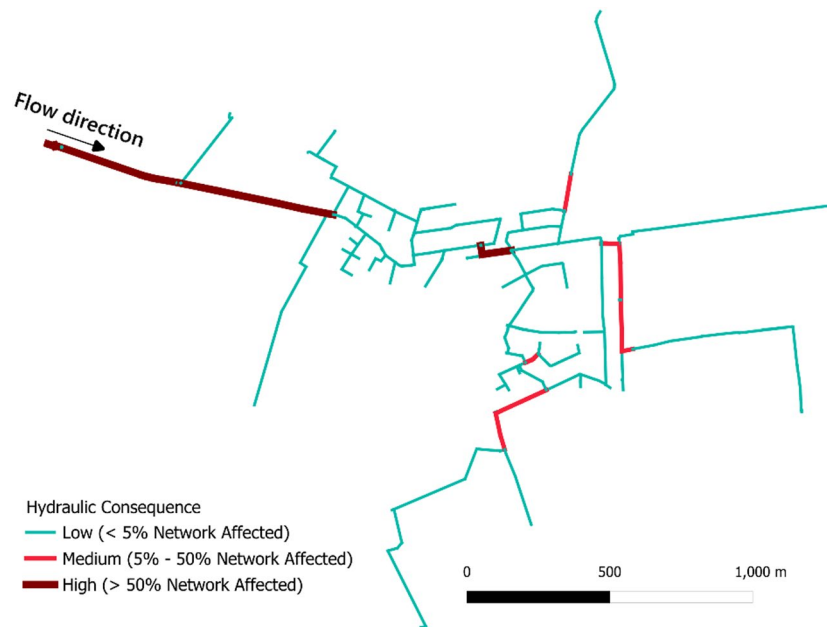
The predicted probabilities were skewed toward smaller values, making direct mapping using fixed probability cut-offs less informative. Therefore, a relative vulnerability (quantile-based) classification was applied so spatial patterns are interpreted in terms of ranked vulnerability rather than absolute probability levels. Pipes were grouped into three tiers based on the predicted-probability distribution: Lower vulnerability ( $\leq 70$ th percentile; bottom 70%), Moderate vulnerability (70th–90th percentile; next 20%), and Higher vulnerability ( $> 90$ th percentile; top 10%). Figure 3 illustrates this spatial mapping. For security reasons, details in the map have been removed and the network configuration have been intentionally distorted.

## Consequence assessment of the case study

### *Hydraulic consequence*

The hydraulic consequence, i.e., impact of a pipe break on the overall system flow and pressure, is shown in Fig. 4. Pipes with high consequence scores are the ones that have a large impact on the hydraulic performance of the network. These pipes are often found in critical locations such as trunk lines feeding large subnetworks, or single-feed pipelines supplying elevated zones. In this study, pipes were categorized into three consequence levels: low (affecting  $< 5\%$  of the network), medium (affecting 5–50% of the network), and high (affecting  $> 50\%$  of the network). These thresholds were selected to provide a clear and interpretable visualization of the network's hydraulic vulnerability. However, these thresholds are adjustable as per utility's own operational standards, risk tolerance, or service-level objectives.

In the case study section, majority of pipe segments exhibited low hydraulic consequence, suggesting that their failure would result in only localized service disruptions. However, a small number of critical mains were associated with high consequence scores, indicating that their failure would lead to significant pressure drops or widespread service interruptions across large areas of the WDN. Notably, one of these high-consequence pipes corresponds to the segment previously evaluated by Odhiambo, et al.<sup>11</sup> for intrusion risks. Other important pipes were mainly located in the central part of the network, branching out from the grid loop. The geography and layout of the network, together with its connectivity to critical demand nodes and the availability of redundant flow paths were found to strongly influence the hydraulic consequence of failures.



**Fig. 4.** Map of the WDN showing pipe segments categorized by the hydraulic disruptions they would cause in the event of a pipe break.

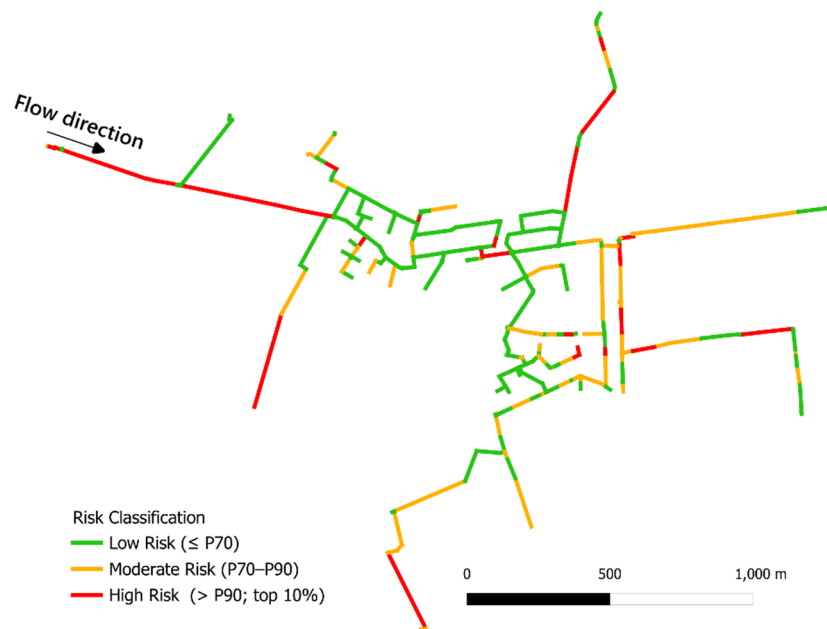
#### Health consequence

The health consequence, obtained after calculating the daily risk of infection, is presented in Fig. 5. Pipe segments were classified into two categories based on the estimated risk: those with a daily infection risk exceeding  $10^{-6}$  and those below the  $10^{-6}$  threshold. The acceptable risk target for daily probability of infection was set at  $10^{-6}$ , to account for short-term periods of heightened risk<sup>68</sup>.

The WDN segments marked in red represent areas that, on average, exceed the acceptable daily risk target throughout the simulation. The spatial pattern of Fig. 5 reveals that certain pipes, often the same ones with high hydraulic impact also rank high in contamination risk. Notably, the highest risks are concentrated in the central section of the WDN and certain branches. This is likely due to a combination of hydraulic conditions, infrastructure vulnerabilities, network topology, and intrusion-related source assumptions (e.g., pathogen concentration and dilution factors). Although other WDNs may behave differently, this approach provides a valuable means of identifying and visualizing pipes with respect to potential health consequences.



**Fig. 5.** Map of the WDN showing pipe segments with a daily risk of infection above  $10^{-6}$  in red.



**Fig. 6.** Map of the WDN showing pipe segments with total risk.

### Risk assessment of the case study

Risk assessment of the pipes in the distribution network was conducted by combining the calibrated probability (P) values, obtained from the RF model, with the corresponding normalized consequence (C) values, derived from both hydraulic and health consequence calculations, using Eq. (2). The calibrated probability values ranged from 0 to 1, while the hydraulic and health consequence values were normalized to a common range of 0 to 1. Thus, the risk scores represent both the probability of failure and the severity of its potential consequences. These values were then visualized spatially using a GIS-based mapping platform, enabling clear, color-coded differentiation of risk levels across the network.

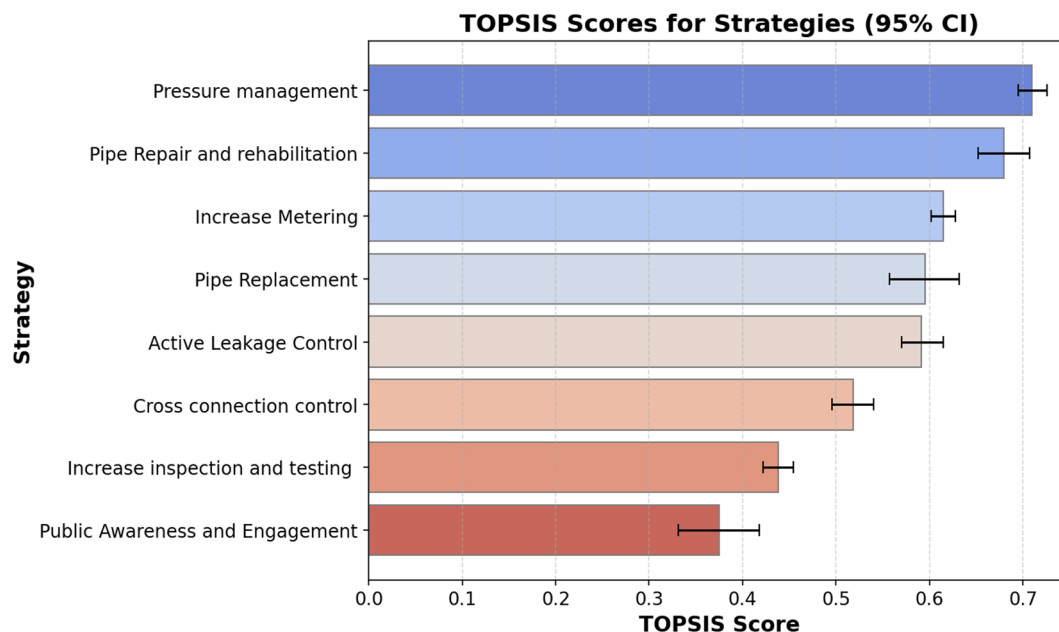
Spatial risk maps facilitate intuitive interpretation, helping utilities identify high-risk zones for targeted maintenance or replacements. Previous studies<sup>69,70</sup> have similarly used spatial mapping in WDNs. Figure 6 presents the spatial distribution of pipe risk classifications within the case study section, highlighting low-risk (bottom 70%), moderate-risk (next 20%), and high-risk (top 10%) segments. The classification was rank-based, based on the distribution of pipe-level risk scores. Accordingly, the three groups corresponded to 6.3 km, 4.1 km, and 3.1 km of pipe, respectively. The resulting map provides a spatial representation of relative risk levels among pipes, enabling comparison of potential vulnerabilities across the network. This can assist water utilities in prioritizing pipe inspection, monitoring, maintenance, and rehabilitation efforts, with pipes in the highest risk tier representing priority candidates for further assessment and targeted intervention, while low-risk pipes can be deprioritized or monitored.

### TOPSIS results

The normalized weights of evaluation criteria, derived from expert input, are presented in Table 7. These weights reflect the relative importance assigned to each criterion within the decision-making process. The results indicate that risk reduction was prioritized most strongly by the experts, highlighting a strong emphasis on minimizing potential health risks and ensuring public safety. Executability and reliability were also highly valued, highlighting the significance of strategies that are feasible and dependable. Cost and social concern received comparatively lower weights but remained non-negligible contributors to the overall decision process. Criteria weights were derived from the expert ratings by first averaging the criterion scores across experts and then normalizing them so that the weights sum to 1. To reflect uncertainty in expert-based judgments, 95%

Criterion	Ranking	Weight (95% CI)
Risk reduction	1	0.266 [0.24, 0.29]
Executability	2	0.201 [0.18, 0.22]
Reliability	3	0.200 [0.17, 0.22]
Cost	4	0.174 [0.14, 0.20]
Social concern	5	0.157 [0.13, 0.18]

**Table 7.** Normalized weights based on expert scores.



**Fig. 7.** Final TOPSIS scores for each alternative (mean expert weights with bootstrap 95% uncertainty intervals).

uncertainty intervals were estimated by bootstrap resampling of experts with replacement and recomputing the normalized criterion weights.

Figure 7 presents the results obtained from the TOPSIS method for strategy prioritization. The ranking highlights ‘Pressure Management’ as the most effective strategy with the highest TOPSIS score (0.71), followed by ‘Pipe Repair and Rehabilitation’ (0.67) and ‘Increase Metering’ (0.61). Strategies such as ‘Pipe Replacement,’ ‘Active Leakage Control’ and ‘Cross-Connection Control’ were ranked in the middle, while ‘Increase inspection and testing’ and ‘Public Awareness and Engagement’ ranked lowest, indicating a lower perceived impact towards health risk reduction.

Expert inputs produced a structured prioritization, underscoring proactive measures over reactive approaches. Pressure management achieved the highest overall score due to its strong performance across multiple criteria. It involves regulating network pressure using methods such as pressure-reducing valves or optimizing pump schedules. This strategy is cost-effective, easily executable, and significantly reduces network-wide risks. This outcome aligns with prior research showing that pressure management can substantially reduce burst frequencies and leakage volumes, while requiring less investment than full pipe replacements<sup>71,72</sup>. Pipe repair and rehabilitation also ranked high, especially regarding risk reduction and reliability, but scored lower in cost and executability, slightly impacting its overall ranking.

Strategies such as active leakage control and increased metering were rated moderately effective due to ease of implementation and lower costs, but they offer less direct risk reduction compared to physical interventions. These results reinforce the value of data-driven, preventive management approaches that proactively mitigate failures. Nevertheless, reactive strategies like pipe replacement remain essential as complementary long-term measures. The findings align with contemporary risk management principles in water supply systems and underscore the importance of balancing technical feasibility, cost-effectiveness, and social considerations to ensure that proposed solutions are both practical and acceptable to stakeholders.

As a robustness check, a weighted sum model (WSM) was also applied using the same criterion weights. In WSM, each strategy receives a score by first normalizing its criterion ratings to a common scale (0–1) and then computing the weighted sum across criteria, so strategies performing well on highly weighted criteria obtain higher overall scores<sup>73</sup>. The WSM-based ranking was compared with the TOPSIS ranking using Spearman’s rank correlation. Spearman’s rank correlation coefficient ( $\rho$ ) measures the degree of agreement between two rankings by correlating the rank positions of the strategies; values close to 1 indicate very similar rankings, values near 0 indicate little agreement, and negative values indicate opposite ordering. The TOPSIS and WSM rankings showed strong agreement ( $\rho = 0.85$ ), indicating that the main prioritization trends were not sensitive to the choice of MCDA aggregation method.

The final TOPSIS rankings are sensitive to the assigned criteria weights. For instance, utilities prioritizing safety might rank pipe replacement higher, whereas those operating under budget constraints may favour cost-effective solutions. The flexibility of TOPSIS accommodates varying organizational goals, risk tolerance, and resource availability, and demonstrates its strength as a robust decision-support method adaptable to diverse utility contexts.

## Conclusions and further work

This study demonstrates an integrated approach to health risk assessment and decision-making within the WSP framework through a case study of a WDN in Sweden. Risk assessment was conducted by integrating ML, QMRA and hydraulic modelling. By applying three ML models (LR, RF, and XGBoost) to estimate the likelihood of pipe breaks and combining it with QMRA and hydraulic modelling, this study effectively identified vulnerable pipes in the WDN. The results were integrated into a GIS framework for the development of a risk map which provided a spatial representation of vulnerability levels within the network.

Expert interviews and stakeholder engagement were utilized to identify key leakage control strategies for health risk reduction, which were subsequently evaluated using MCDA. Results from the TOPSIS method aligned well with WSP goals, prioritizing pressure management and pipe rehabilitation. These findings highlight the significance of investing in infrastructure maintenance and operational efficiency to mitigate health risks and improve network reliability. Employing MCDA also facilitated stronger stakeholder collaboration for risk management, establishing a robust foundation for informed decision-making regarding health risks from pipe failures and ensuring sustainable water supply systems.

The integrated risk results can support different intervention pathways depending on whether risk is driven mainly by probability of failure or consequence of failure. Pipes with high probability but low consequence may be addressed through preventive actions such as inspection, leak detection, pressure monitoring, valve exercising, or targeted maintenance. In contrast, pipes with lower probability but higher consequence may require measures aimed at reducing impacts, such as improved redundancy, isolation capacity, protection of critical customers, or contingency planning. Pipes ranking high in both dimensions may warrant more intensive interventions, including rehabilitation or replacement.

Several areas were identified for further research. Since the case study network is relatively small, the reported performance metrics and prioritization results should not be assumed to generalize beyond the studied context. At the same time, this setting reflects a common reality for many small and medium utilities with relatively few failures and limited data access, and the case study demonstrates that the proposed WSP workflow can still be implemented in practice by linking failure likelihood, consequence modelling, and decision support under realistic data limitations. Expanding the dataset (e.g., additional years, larger service areas, or multiple utilities), together with more systematic model optimization, would strengthen predictive accuracy, enable stronger external validation, and support more robust inferences.

Further improvements are also possible on the predictive modelling side. Incorporating time-dependent variables would enable more accurate forecasting of pipe failures, which is important for proactive planning. While the evaluated ML models performed reasonably, future studies could extend the comparison to additional model families (e.g., neural networks) and adopt robustness-oriented workflows such as nested cross-validation, together with more systematic hyperparameter optimization (e.g., Bayesian optimization). Model ensembling and explicit cost-sensitive learning could also be explored to better reflect utility-specific inspection and rehabilitation constraints and to align model outputs with operational decision-making. In practice, such outputs could support a simple decision workflow in which higher-risk pipes are first identified, followed by selection of the most appropriate intervention strategy using the TOPSIS rankings under local utility constraints and priorities. Additionally, the 'Impact Score' feature, despite capturing valuable network interdependencies, requires extensive hydraulic simulations. Future research could explore optimized simulation techniques to enhance computational scalability.

Since pipe-break datasets are typically highly imbalanced, random oversampling was applied in this study to increase the representation of the minority class due to its simplicity. However, oversampling duplicates minority observations and may increase overfitting risk. Future studies could compare alternative imbalance strategies (e.g., class-weight adjustment or synthetic oversampling such as SMOTE) and evaluate their effect on both discrimination and probability quality, particularly when larger datasets are available.

On the consequence and decision-support side, the analysis can be extended to include additional consequence dimensions such as repair costs, service downtime, customer impacts, or proximity to critical infrastructure, providing a more comprehensive basis for pipe-level risk assessment. The MCDA component can also be strengthened by combining expert judgement with available quantitative indicators where possible, periodically updating stakeholder inputs to reflect changing priorities, and validating prioritization outcomes against implemented interventions and observed performance improvements.

The proposed workflow spans multiple sources of uncertainty, including uncertainty in failure likelihood estimates, hydraulic parameters, QMRA inputs, and expert judgement used in MCDA, which can introduce some subjectivity in the criteria weighting. In this study, uncertainty was addressed at two practical decision points: uncertainty bounds were reported for ML discrimination performance, and bootstrap-based uncertainty intervals were computed for expert-derived criteria weights and resulting TOPSIS scores. Building on this, future work can move from component-level uncertainty reporting toward end-to-end uncertainty propagation, quantifying how uncertainty in probability and consequence jointly affects final risk estimates and strategy rankings. Similarly, alternative risk formulations can be explored, to better represent nonlinear effects, cascading failures, and multiple interacting scenarios (e.g., scenario-based analyses, fault-/event-tree logic, or Monte Carlo propagation). Finally, the GIS mapping component was constrained by security requirements, which required removal or distortion of certain cartographic elements. This limits quantitative spatial interpretation, including detailed assessment of network density and the spatial concentration of higher-risk areas. However, the risk map remains useful for decision support by showing relative vulnerability patterns and highlighting priority areas for inspection and intervention.

In conclusion, this study provides a practical demonstration of integrating predictive modelling, consequence evaluation, and structured decision-making within the WSP framework. The application of this approach in a real-world case study illustrates how utilities with limited data can identify vulnerable assets, assess potential

health consequences, and prioritize risk mitigation strategies. This approach is adaptable to different network sizes, data availability, and operational contexts, making it useful for a wide range of water utilities. By combining technical analysis with stakeholder input, it strengthens risk-based planning and supports more informed decision-making in water supply systems.

### Data availability

The data that support the findings of this study are available from the collaborating water utility in Skåne, Sweden, but restrictions apply to the availability of these data, which were used under licence for the current study and are therefore not publicly available. Data are however available from the corresponding author upon reasonable request and with permission of the water utility.

Received: 4 June 2025; Accepted: 5 May 2026

Published online: 15 May 2026

### References

- Jafari, S. M., Zahiri, A. R. & Hadad, B. Mohammad Rezapour Tabari, M. A hybrid of six soft models based on ANFIS for pipe failure rate forecasting and uncertainty analysis: a case study of Gorgan city water distribution network. *Soft Comput.* **25**, 7459–7478. <https://doi.org/10.1007/s00500-021-05706-4> (2021).
- Adedeji, K., Hamam, Y., Abe, B. & Abu-Mahfouz, A. Leakage Detection and Estimation Algorithm for Loss Reduction in Water Piping Networks. *Water* **9**, 773. <https://doi.org/10.3390/w9100773> (2017).
- Liemberger, R. & Wyatt, A. Quantifying the global non-revenue water problem. *Water Supply*. **19**, 831–837. <https://doi.org/10.2166/ws.2018.129> (2018).
- Sörensen, J., Nilsson, E., Bjarke, M., Giertz, T. & Nilsson, D. Sanning och konsekvens för svenska vattenledningar | Truth and dare for Swedish water pipes. (2021).
- Besner, M. C., Prévost, M. & Regli, S. Assessing the public health risk of microbial intrusion events in distribution systems: Conceptual model, available data, and challenges. *Water Res.* **45**, 961–979 (2011).
- Raspati, G. S. et al. A Risk-Based Approach in Rehabilitation of Water Distribution Networks. *Int. J. Environ. Res. Public Health*. **19**, 1594. <https://doi.org/10.3390/ijerph19031594> (2022).
- World Health Organization. *Water safety in distribution systems* (World Health Organization, 2014).
- Roeger, A. & Tavares, A. F. Do Governance Arrangements Affect the Voluntary Adoption of Water Safety Plans? An Empirical Study of Water Utilities in Portugal. *Water Resour. Manage.* **34**, 1757–1772. <https://doi.org/10.1007/s11269-020-02527-2> (2020).
- Gunnarsdottir, M. J., Gardarsson, S. M., Elliott, M., Sigmundsdottir, G. & Bartram, J. Benefits of water safety plans: microbiology, compliance, and public health. *Environ. Sci. Technol.* **46**, 7782–7789 (2012).
- Taiwo, R., Zayed, T. & Ben Seghier, M. E. A. Integrated intelligent models for predicting water pipe failure probability. *Alexandria Eng. J.* **86**, 243–257. <https://doi.org/10.1016/j.aej.2023.11.047> (2024).
- Odhiambo, M., Viñas, V., Sokolova, E. & Pettersson, T. J. R. Health risks due to intrusion into the drinking water distribution network: hydraulic modelling and quantitative microbial risk assessment. *Environ. Science: Water Res. Technol.* **9**, 1701–1716. <https://doi.org/10.1039/D2EW00720G> (2023).
- Park, H., Kim, K., Hyung, J., Kim, T. & Koo, J. Decision-making for the hazard ranking of water distribution networks using the TOPSIS method. *Water Supply*. **23**, 715–726. <https://doi.org/10.2166/ws.2023.013> (2023).
- Liu, W., Wang, B. & Song, Z. Failure Prediction of Municipal Water Pipes Using Machine Learning Algorithms. *Water Resour. Manage.* **36**, 1271–1285. <https://doi.org/10.1007/s11269-022-03080-w> (2022).
- Asadi, Y. Employing machine learning in water infrastructure management: predicting pipeline failures for improved maintenance and sustainable operations. *Industrial Artif. Intell.* **2**, 8. <https://doi.org/10.1007/s44244-024-00022-w> (2024).
- Haas, C. N., Rose, J. B. & Gerba, C. P. *Quantitative microbial risk assessment* (Wiley, 2014).
- Rossman, L. A. et al. EPANET 2.2 User Manual; Water Infrastructure Division. *Center for Environmental Solutions and Emergency Response* (2020).
- Majumder, M. & Saha, A. K. in *Feasibility Model of Solar Energy Plants by ANN and MCDM Techniques*. 9–12 (eds Majumder, M., Apu, K. & Saha) (Springer, 2016).
- Hwang, C. L. & Yoon, K. in *Multiple Attribute Decision Making: Methods and Applications A State-of-the-Art Survey* (eds Ching-Lai Hwang & Kwangsun Yoon) 58–191 Springer, (1981).
- Rausand, M. *Risk assessment: theory, methods, and applications*. **115** (Wiley, 2013).
- Robles-Velasco, A., Cortés, P., Muñuzuri, J. & Onieva, L. Prediction of pipe failures in water supply networks using logistic regression and support vector classification. *Reliab. Eng. Syst. Saf.* **196**, 106754. <https://doi.org/10.1016/j.res.2019.106754> (2020).
- Francis, R. A., Guikema, S. D. & Henneman, L. Bayesian Belief Networks for predicting drinking water distribution system pipe breaks. *Reliab. Eng. Syst. Saf.* **130**, 1–11. <https://doi.org/10.1016/j.res.2014.04.024> (2014).
- Barton, N. A., Farewell, T. S., Hallett, S. H. & Acland, T. F. Improving pipe failure predictions: Factors affecting pipe failure in drinking water networks. *Water Res.* **164**, 114926. <https://doi.org/10.1016/j.watres.2019.114926> (2019).
- Makar, J. M. A preliminary analysis of failures in grey cast iron water pipes. *Eng. Fail. Anal.* **7**, 43–53. [https://doi.org/10.1016/S1350-6307\(99\)00005-9](https://doi.org/10.1016/S1350-6307(99)00005-9) (2000).
- Røstum, J. *Statistical modelling of pipe failures in water networks* (Fakultet for ingeniørvitenskap og teknologi, 2000).
- Fan, X., Wang, X., Zhang, X. & Yu, X. Machine learning based water pipe failure prediction: The effects of engineering, geology, climate and socio-economic factors. *Reliab. Eng. Syst. Saf.* **219**, 108185. <https://doi.org/10.1016/j.res.2021.108185> (2022).
- Wilson, D., Fillion, Y. & Moore, I. State-of-the-art review of water pipe failure prediction models and applicability to large-diameter mains. *Urban Water J.* **14**, 173–184. <https://doi.org/10.1080/1573062X.2015.1080848> (2017).
- Kleiner, Y. & Rajani, B. Comprehensive review of structural deterioration of water mains: statistical models. *Urban Water*. **3**, 131–150. [https://doi.org/10.1016/S1462-0758\(01\)00033-4](https://doi.org/10.1016/S1462-0758(01)00033-4) (2001).
- Giraldo-González, M. M. & Rodríguez, J. P. Comparison of Statistical and Machine Learning Models for Pipe Failure Modeling in Water Distribution Networks. *Water* **12**, 1153. <https://doi.org/10.3390/w12041153> (2020).
- Forero-Ortiz, E. et al. Models and explanatory variables in modelling failure for drinking water pipes to support asset management: a mixed literature review. *Appl. Water Sci.* **13**, 210. <https://doi.org/10.1007/s13201-023-02013-1> (2023).
- Dawood, T., Elwakil, E. & Mayol Novoa, H. Fernando Gárate Delgado, J. Pressure data-driven model for failure prediction of PVC pipelines. *Eng. Fail. Anal.* **116**, 104769. <https://doi.org/10.1016/j.engfailanal.2020.104769> (2020).
- Kunz, J. M. Surveillance of Waterborne Disease Outbreaks Associated with Drinking Water—United States, 2015–2020. *MMWR Surveillance Summaries* **73** (2024).
- Guzman-Herrador, B. et al. Waterborne outbreaks in the Nordic countries, 1998 to 2012. *Eurosurveillance* **20** (2015).
- Muddassir, M., Zayed, T., Taiwo, R. & Ben Seghier, M. E. A. Advancing the analysis of water pipe failures: a probabilistic framework for identifying significant factors. *Sci. Rep.* **14**, 19218. <https://doi.org/10.1038/s41598-024-69855-w> (2024).

34. Nygård, K. et al. Breaks and maintenance work in the water distribution systems and gastrointestinal illness: a cohort study. *Int. J. Epidemiol.* **36**, 873–880 (2007).
35. Sæve-Söderbergh, M., Bylund, J., Malm, A., Simonsson, M. & Toljander, J. Gastrointestinal illness linked to incidents in drinking water distribution networks in Sweden. *Water Res.* **122**, 503–511. <https://doi.org/10.1016/j.watres.2017.06.013> (2017).
36. Viñas, V., Sokolova, E., Malm, A., Bergstedt, O. & Pettersson, T. J. R. Cross-connections in drinking water distribution networks: Quantitative microbial risk assessment in combination with fault tree analysis and hydraulic modelling. *Sci. Total Environ.* **831**, 154874. <https://doi.org/10.1016/j.scitotenv.2022.154874> (2022).
37. Zolghadr-Asli, B., Bozorg-Haddad, O. & van Cauwenbergh, N. in *Essential Tools for Water Resources Analysis, Planning, and Management* (ed Omid Bozorg-Haddad) 305–322 Springer, (2021).
38. Ngunbane, Z., Bergion, V., Dzwauro, B., Stenström, T. A. & Sokolova, E. Multi-criteria decision analysis framework for engaging stakeholders in river pollution risk management. *Sci. Rep.* **14** <https://doi.org/10.1038/s41598-024-57739-y> (2024).
39. Farmani, R., Walters, G. & Savic, D. Evolutionary multi-objective optimization of the design and operation of water distribution network: total cost vs. reliability vs. water quality. *J. Hydroinformatics.* **8**, 165–179 (2006).
40. Aboutaleb, M., Bozorg-Haddad, O. & Loáiciga, H. A. Multiobjective design of water-quality monitoring networks in river-reservoir systems. *J. Environ. Eng.* **143**, 04016070 (2017).
41. Mutikanga, H. E., Sharma, S. K. & Vairavamoorthy, K. Multi-criteria Decision Analysis: A Strategic Planning Tool for Water Loss Management. *Water Resour. Manage.* **25**, 3947–3969. <https://doi.org/10.1007/s11269-011-9896-9> (2011).
42. Ghandi, M. & Roozbahani, A. Risk Management of Drinking Water Supply in Critical Conditions Using Fuzzy PROMETHEE V Technique. *Water Resour. Manage.* **34**, 595–615. <https://doi.org/10.1007/s11269-019-02460-z> (2020).
43. Salehi, S., Jalili Ghazizadeh, M. & Tabesh, M. A comprehensive criteria-based multi-attribute decision-making model for rehabilitation of water distribution systems. *Struct. Infrastruct. Eng.* **14**, 743–765. <https://doi.org/10.1080/15732479.2017.1359633> (2018).
44. Scholten, L., Scheidegger, A., Reichert, P., Mauer, M. & Lienert, J. Strategic rehabilitation planning of piped water networks using multi-criteria decision analysis. *Water Res.* **49**, 124–143. <https://doi.org/10.1016/j.watres.2013.11.017> (2014).
45. Omkarprasad, S. V. & Sushil, K. Analytic hierarchy process: An overview of applications. *Eur. J. Oper. Res.* **169**, 1–29. <https://doi.org/10.1016/j.ejor.2004.04.028> (2006).
46. Govindan, K. & Jepsen, M. B. ELECTRE: A comprehensive literature review on methodologies and applications. *Eur. J. Oper. Res.* **250**, 1–29 (2016).
47. James, G., Witten, D., Hastie, T. & Tibshirani, R. *An introduction to statistical learning.* **112** (Springer, 2013).
48. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32. <https://doi.org/10.1023/A:1010933404324> (2001).
49. Chen, T. & Guestrin, C. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 Association for Computing Machinery, San Francisco, California, USA, (2016).
50. Aslani, B., Mohebbi, S. & Axthelm, H. Predictive analytics for water main breaks using spatiotemporal data. *Urban Water J.* **18**, 433–448. <https://doi.org/10.1080/1573062X.2021.1893363> (2021).
51. Beig Zali, R., Latifi, M., Javadi, A. A. & Farmani, R. Semisupervised Clustering Approach for Pipe Failure Prediction with Imbalanced Data Set. *J. Water Resour. Plan. Manag.* **150**, 04023078. <https://doi.org/10.1061/JWRMD5.WRENG-6263> (2024).
52. Daulat, S., Rokstad, M. M., Bruaset, S., Langeveld, J. & Tschekner-Gratl, F. Evaluating the generalizability and transferability of water distribution deterioration models. *Reliab. Eng. Syst. Saf.* **241**, 109611. <https://doi.org/10.1016/j.res.2023.109611> (2024).
53. Bishop, C. M. & Nasrabadi, N. M. *Pattern recognition and machine learning.* **4** (Springer, 2006).
54. Latifi, M., Beig Zali, R., Javadi, A. A. & Farmani, R. Customised-sampling approach for pipe failure prediction in water distribution networks. *Sci. Rep.* **14**, 18224. <https://doi.org/10.1038/s41598-024-69109-9> (2024).
55. de Werner, V., Schneider Aranda, J. A., dos Santos Costa, R., da Silva Pereira, P. R. & Victória Barbosa, J. L. Imbalanced data preprocessing techniques for machine learning: a systematic mapping study. *Knowl. Inf. Syst.* **65**, 31–57. <https://doi.org/10.1007/s10115-022-01772-8> (2023).
56. Stehman, S. V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* **62**, 77–89 (1997).
57. Klise, K. A., Bynum, M., Moriarty, D. & Murray, R. A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study. *Environ. Model. Softw.* **95**, 420–431. <https://doi.org/10.1016/j.envsoft.2017.06.022> (2017).
58. Wagner, J. M., Shamir, U. & Marks, D. H. Water distribution reliability: Simulation methods. *J. Water Resour. Plan. Manag.* **114**, 276–294 (1988).
59. Hewitt, J., Leonard, M., Greening, G. E. & Lewis, G. D. Influence of wastewater treatment process and the population size on human virus profiles in wastewater. *Water Res.* **45**, 6267–6276 (2011).
60. Metcalf & Eddy, I. *Wastewater Engineering: Treatment and Reuse* (McGraw-Hill Education, 2003).
61. Abrahamsson, J. L., Ansker, J. & Heinicke, G. *MRA – a model for Swedish water works (in Swedish)* (Swedish Water and Wastewater Association, 2009).
62. Teunis, P. et al. A reconsideration of the *Campylobacter* dose-response relation. *Epidemiol. Infect.* **133**, 583–592. <https://doi.org/10.1017/s0950268805003912> (2005).
63. Teunis, P. F. M. et al. Norwalk virus: How infectious is it? *J. Med. Virol.* **80**, 1468–1476. <https://doi.org/10.1002/jmv.21237> (2008).
64. Teunis, P. F. M., Chappell, C. L. & Okhuysen, P. C. Cryptosporidium Dose-Response Studies: Variation Between Hosts. *Risk Anal.* **22**, 475–485. <https://doi.org/10.1111/0272-4332.00046> (2002).
65. Taiwo, R., Shaban, I. A. & Zayed, T. Development of sustainable water infrastructure: A proper understanding of water pipe failure. *J. Clean. Prod.* **398**, 136653. <https://doi.org/10.1016/j.jclepro.2023.136653> (2023).
66. Winkler, D., Haltmeier, M., Kleidorfer, M., Rauch, W. & Tschekner-Gratl, F. Pipe failure modelling for water distribution networks using boosted decision trees. *Struct. Infrastruct. Eng.* **14**, 1402–1411. <https://doi.org/10.1080/15732479.2018.1443145> (2018).
67. Konstantinou, C. & Stoianov, I. A comparative study of statistical and machine learning methods to infer causes of pipe breaks in water supply networks. *Urban Water J.* **17**, 534–548. <https://doi.org/10.1080/1573062X.2020.1800758> (2020).
68. Signor, R. & Ashbolt, N. Comparing probabilistic microbial risk assessments for drinking water against daily rather than annualised infection probability targets. *J. Water Health.* **7**, 535–543 (2009).
69. Nunes, R., Arraut, E. & Pimentel, M. Risk Assessment Model for the Renewal of Water Distribution Networks: A Practical Approach. *Water* **15**, 1509. <https://doi.org/10.3390/w15081509> (2023).
70. Tabesh, M. & Saber, H. A Prioritization Model for Rehabilitation of Water Distribution Networks Using GIS. *Water Resour. Manage.* **26**, 225–241. <https://doi.org/10.1007/s11269-011-9914-y> (2012).
71. Jara-Arriagada, C. & Stoianov, I. Pipe breaks and estimating the impact of pressure control in water supply networks. *Reliab. Eng. Syst. Saf.* **210**, 107525 (2021).
72. Tian, Y. et al. Optimization of pressure management in water distribution systems based on pressure-reducing valve control: Evaluation and case study. *Sustainability* **15**, 11086 (2023).
73. Triantaphyllou, E. in *Multi-criteria decision making methods: A comparative study* 5–21 (Springer, 2000).

## Acknowledgements

The authors are grateful to the water utility staff at the municipality for providing information and data for the modelling.

### Author contributions

All authors contributed to the study conception and design. Uchit Sangroula contributed to conceptualization, methodology design, original draft preparation, investigation, data curation, and formal analysis. Victor Viñas and Michael Odhiambo contributed to conceptualization, data curation, methodology, writing, and review. Thomas J.R. Pettersson contributed to the initiation of the study, conceptualization, supervision, and the review and editing of the manuscript.

### Funding

Open access funding provided by Chalmers University of Technology. This work was financially supported by the Swedish Research Council for Sustainable Development FORMAS project AquaClim: Building a Better Climate with Water Research (project number 2022 – 01900).

### Declarations

#### Competing interests

The authors declare no competing interests.

#### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-026-52465-z>.

**Correspondence** and requests for materials should be addressed to U.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026