Models and explanatory variables in modelling failure for drinking water pipes to support asset management: a mixed literature review

Edwar Forero-Ortiz[1,2] · Eduardo Martinez-Gomariz[1,2] · Marti Sanchez-Juny[2] · Jaume Cardus Gonzalez[1] · Fernando Cucchietti[3] · Ferran Baque Viader[4] · Miquel Sarrias Monton[5]

Abstract
There is an increasing demand to enhance infrastructure asset management within the drinking water sector. A key factor for achieving this is improving the accuracy of pipe failure prediction models. Machine learning-based models have emerged as a powerful tool in enhancing the predictive capabilities of water distribution network models. Extensive research has been conducted to explore the role of explanatory variables in optimizing model outputs. However, the underlying mechanisms of incorporating explanatory variable data into the models still need to be better understood. This review aims to expand our understanding of explanatory variables and their relationship with existing models through a comprehensive investigation of the explanatory variables employed in models over the past 15 years. The review underscores the importance of obtaining a substantial and reliable dataset directly from Water Utilities databases. Only with a sizeable dataset containing high-quality data can we better understand how all the variables interact, a crucial prerequisite before assessing the performance of pipe failure rate prediction models.

✉ Edwar Forero-Ortiz
eaforero@aiguesdebarcelona.cat;
edwar.andres.forero@upc.edu

1 Aigües de Barcelona, Empresa Metropolitana de la Gestió del Cicle Integral de l'Aigua, C/General Batet 1-7, 08028 Barcelona, Spain

2 Flumen Research Institute, Universitat Politècnica de Catalunya - Centre Internacional de Mètodes Numèrics en Enginyeria, Barcelona, Spain

3 Barcelona Supercomputing Center, C/Jordi Girona 31, 08034 Barcelona, Spain

4 AMB, Àrea Metropolitana de Barcelona, Servei de Supervisió de Concessions, Direcció de Serveis del Cicle de l'Aigua, C/62 16-18, edifici A - Zona Franca, 08040 Barcelona, Spain

5 Water Technology Centre, Cornellà de Llobregat, Spain

Introduction

Asset management and its relationship with water infrastructure modelling

Significant investments are necessary for the construction, management, and maintenance of water supply and distribution network (WSDN) infrastructure to ensure access to drinking water that meets the required standards in terms of quality and quantity for the population. These investments introduce an additional layer of complexity, encompassing various technical, economic, social, and environmental aspects within the system (Sitzenfrei et al. 2020; Kerwin and Adey 2021; Meijer et al. 2021). The foreseen deadlines for the components' end-of-life in WSDNs underscore the imperative for enhanced planning and efficient resource allocation to mitigate service level deterioration and infrastructure conditions. These concerns carry significant global implications, threatening the overall quality of provided services (Zamenian et al. 2017; Bello et al. 2019; Curt et al. 2019; Vieira et al. 2020).

The repercussions of WSDN deterioration processes, including pipe breaks, water losses, substandard service delivery, compromised water supply quality, and escalating operational and maintenance expenses, are well-documented (Bello et al. 2019). Water distribution pipes play a crucial role in water supply and distribution systems, and their failures have significant financial, social, and environmental ramifications (Fan et al. 2021). The pipe replacement rate by drinking water companies ranges from 0.5 to 4.8% per year, exemplified by the average annual renewal of 0.67% of the network length in the case of France (Office Français de la Biodiversité 2022). In comparison, the USA exhibits an average rate ranging from 1 to 4.8%, representing a one-time renewal rate. Consequently, it would take approximately 20 to 200 years to renew the current water infrastructures worldwide fully. Table 1 summarizes the infrastructure and renewal rates, water losses, leakage, and non-revenue water in various countries' WSDNs. The deterioration of pipelines significantly impacts WSDNs globally, evident in metrics such as water losses and the associated costs, insufficient network renewal results in deteriorating pipeline conditions and increased system failures.

The challenges posed by water infrastructure necessitate a resilient and interdisciplinary approach to implementing an asset management system. Such a system encompasses a cohesive network of elements within an organization, defining the asset management policy, establishing asset management objectives, and formulating the necessary processes to achieve these objectives (ISO/TC 251 2014).

Asset management systems are intricately linked to water infrastructure management, operating within the framework of Infrastructure Asset Management (IAM). IAM encompasses essential activities geared towards optimizing service delivery, including inventorying, monitoring, maintenance, and renovation. This holistic approach seamlessly integrates engineering and management sciences, aligning technical aspects with usage, perception, and value considerations. Such integration enhances informed decision-making and facilitates the development of efficient management strategies. By intertwining engineering sciences with management disciplines, IAM is pivotal for Water Utilities in ensuring the long-term maintenance and adaptability of water infrastructure, effectively addressing ageing and potential obsolescence (Le Gat et al. 2023).

Modelling and optimization in the operation of structures, coupled with probabilistic modelling of structural deterioration (performance failure), represent significant and forward-looking trends as underscored by Le Gat et al. (2023), shaping the strategic objectives of Water Infrastructure Asset Management (WIAM). These trends align with the findings of other prominent authors who have extensively researched IAM. Notable researchers in recent years include Ugarelli and Sægrov (2022), Okwori et al. (2021), Pathirana et al. (2021), El-Diraby (2021), Mazumder et al. (2021), Beuken et al. (2020), Curt et al. (2019), and Carriço et al. (2020), among others in recent years. These researchers note the adaptation of WIAM frameworks to embrace digitization and the implementation of modelling in diverse domains,

Table 1 Status of WSDNs at the national level in terms of infrastructure and renewal rates, water losses and leakage, and non-revenue water

Year	Country	Status of WSDN	Source
2021	USA	By 2019, Water utilities replaced between 1 and 4.8% of their pipelines per year on average, a replacement rate that matches water pipes' lifecycle	ASCE (2022)
2021	Portugal	28.7% of the total water in the distribution system is non-revenue water, with actual water losses of 174 million m^3 of water/year	ERSAR (2021)
2020	Spain	The overall network renewal rate is 0.43%, and over 27% of the pipeline is more than 40 years old	AEAS (2021)
2020	The UK	Almost 1095 million m^3 of water (20% of the total water supply) are lost daily due to leaks	National Audit Office (2020)
2020	USA, Canada, and Mexico	In North America, the total cost of water losses due to pipe breaks is estimated at USD 3.8 billion per year	Snider and McBean (2020)
2020	Canada	There are over 6000 km of water pipelines in Toronto, 13% are between 80 and 100 years old, and 11% are over 100 years old	El-Diraby (2021)
2019	China	According to the National Bureau of Statistics of China, in 2019, some cities had water leakage rates higher than 40%	Liu et al. (2022)
2019	Canada	17,788 km (9.6%) of the pipelines were in poor/deplorable condition, and 32,641 km (17.7%) were in fair condition	CIRC (2019)
2019	USA	Drinking water systems currently lose 7.95×10^3 million m^3 of non-revenue water loss annually, losing an estimated USD 7.6 billion of treated water in 2019 due to leaks, with more than 240,000 pipeline breaks	ASCE (2022)
2018	The USA and Canada	Between 2012 and 2018, overall water pipelines break rates increased by 27%, from 11.0 to 14.0 breaks/161 km/year	Folkman (2018)

as expounded in detail by Okwori et al. (2021) and Kerwin and Adey (2021). As a result, WIAM is evolving and paving the way for sustainable water infrastructure practices by incorporating asset modelling.

Modelling in water infrastructure asset management: significance and diverse outputs at various levels

Models play a crucial role in WIAM as they enable a systematic approach to understanding and optimizing the operation of water infrastructure systems. They provide valuable insights into the complex dynamics of these systems, helping to enhance decision-making processes and improve overall performance (Garzón et al. 2022).

Modelling is vital in comprehending the complex dynamics of water distribution systems, involving the intricate interplay between water, the environment, and the ageing infrastructure of WSDN. Employing data-driven models facilitates the extraction of meaningful insights from data patterns, capturing variable relationships without predetermined mechanisms or interactions. In the context of WIAM-focused WSDN modelling, historical datasets of leaks and pipe failures serve as the foundation to explore factors influencing pipe lifespan. These influential factors, termed explanatory variables, constitute this paper's primary focus of investigation.

WIAM relies on models that perform several critical roles, such as prediction; models can estimate the performance of water infrastructure assets over time. For instance, the modelling approach in this study considers age, material, environmental and social conditions, among other factors, to predict the likelihood of asset failures, enabling the implementation of proactive maintenance and replacement strategies and mitigating potential issues. Optimization, using models, can enhance the operation and maintenance of water infrastructure assets (Ulusoy et al. 2021). Models can identify the most efficient asset management strategies by assessing cost, resource availability, sustainability metrics, and system performance.

Risk Assessment, through modelling, allows for evaluating potential risks related to various asset management decisions. These models can measure the potential outcomes of asset failures and pinpoint areas of significant vulnerability, thereby directing efforts towards mitigating risk (Ugarelli and Sægrov 2022). Models facilitate scenario analysis, allowing decision-makers to explore different "what-if" scenarios and their potential impacts, helping to evaluate the effectiveness of different strategies and identify the best course of action (Rulleau et al. 2020).

Regarding the types of output provided by models at different levels, it depends on the scope and complexity of the model and the level of the water infrastructure system being analysed. At the network level, models may yield outputs related to the overall system performance, such as predicting the failure rate, estimating water loss, and assessing energy consumption; also offering valuable insights into network-wide asset conditions and identifying critical assets require immediate attention (Fan et al. 2023).

At the asset level, models can provide predictions specific to individual components, such as pipelines or pumps, which may include the probability of failure for each asset, estimations of remaining useful life, and optimal maintenance schedules. At the strategic level, models can offer outputs that guide long-term planning and investment decisions, identifying areas requiring infrastructure upgrades, assessing the impact of various investment scenarios, and optimizing capital allocation (Mohammadi and Amador Jimenez 2022).

Models also contribute outputs at the tactical level, facilitating day-to-day decision-making, influencing the prioritization of maintenance activities, optimizing inspection schedules, and identifying short-term risk mitigation measures (Alegre et al. 2013). Thus, models play a crucial role in WIAM by predicting asset performance, optimizing operations, assessing risks, and enabling informed decision-making (Le Gat et al. 2023). The nature of the outputs they provide varies depending on the level of analysis and the specific objectives of the asset management process. Notably, in the context of this study, the models estimating Drinking Water Pipe Failure (DWPF) outputs for WSDN systems offer outcomes that address all identified levels.

Bridging the divide: identifying gaps and addressing the necessity of a review on explanatory variables in drinking water pipe failure models

Numerous review articles have comprehensively summarized and synthesized various facets of water pipe failure modelling. Notably, Rostum (2000) provides an overview of developments up to 2000, including the seminal thesis by Eisenbeis (1994). The works of Pelletier (2000; Mailhot et al. 2000), Mailhot et al. (2003), among others, have made significant contributions to this line of research, which was initially initiated by Shamir and Howard (1979) and further expanded upon by Kettler and Goulter (1985) and Andreou et al. (1987) with a primary focus on developing methodologies for enhancing the maintenance of long-established WSDNs. To provide a comprehensive understanding of the broader aspects of water pipe failure modelling, Table 2 presents notable publications that aim to encapsulate the field without explicitly delving into the definition and analysis of explanatory variables in failure processes.

Three distinct publications have examined the factors influencing failure prediction models within water infrastructure. One study explicitly investigates water quality

Table 2 Previous reviews in the field of modelling for predicting DWPFs

Source	Types of models considered	Type of prediction/nature of the output	Described items for each model	Final recommendations
St. Clair and Sinha (2012)	Deterministic, statistical, probabilistic, and advanced mathematical models consist of artificial neural networks (ANN), fuzzy logic and heuristic	Remaining service life, prioritizing replacement, risks of a pipe burst or failure, strength, failure rates, optimal replacement, break rates, lifetime, residual life, deterioration rates, condition rating, pipe breaks, and vulnerability rates	Data requirements for each model and publication subtype	An in-depth review and recommendations for enhancing WIAM Practices
Nishiyama and Filion (2013)	Statistical deterministic models, statistical probabilistic risk models, fuzzy-based techniques, and artificial neural networks	Number of breaks, annual break rate, failure frequency, probability of observing breaks, the occurrence of failure, optimal replacement time, failure intensity, the time between consecutive breaks	Discussion, general model equations, covariates used in equations, available model procedure, statistical tests for model validation, model limitations	Feasible future implementations and improvements to every analysed model
Scheidegger et al. (2013)	The review focused on Statistical methods: Weibull distributions, semi-Markov process, exponential, non-homogeneous Poisson process, Yule process	Failure rate, maximum failure likelihood	Model structure, calibration, calibration data, model applications, further developments, predictions, software	It presents a summary of the statistical models' properties as of the date of the study, including a comparison among types of deterioration models
Large (2015)	Deterministic and probabilistic pipeline degradation prediction models, risk/impact models, and economic and decision models	Number of breaks, days without failures, pipeline age in the first breakdown, remaining service life, cumulated failures, predicted water interruptions, critical hydraulic index, expected volume of unsatisfied demand in a water utility, and reliability index	Model description, hypotheses and formulas, advantages and disadvantages of the model, and synthesis	None
Wilson et al. (2017)	The review focused on large-diameter mains. Classifications: Physical and Statistical. Statistical: Proportional hazards, Markov processes, Hierarchical process	Remaining service life, corrosion pit depth, structural capacity, time to failure, failure rate, probability of failure, the likelihood of failure, the hazard function, and life-risk curve	Model description, formulae, and critique	The mechanisms underlying pipe failure are intricate and require a comprehensive understanding. Notably, larger-diameter pipes exhibit a lower failure rate compared to smaller-diameter pipes, posing challenges in accurately predicting pipe failures
(Gao 2017)	Deterministic, probabilistic, stochastic, and artificial intelligence and machine learning methods	Remaining useful life, failure rate, probability of failure, number of breaks	Model description, approach, considerations, and data requirements	This MSc Thesis presents two pivotal preliminary findings: a comprehensive review of contributing factors for pipe failures and conclusive insights on these explanatory variables
Dawood et al. (2020b)	Statistical methods, artificial neural network-based models, fuzzy logic-based models	Likelihood of pipe failures, failure rate, number of pipe failures, deterioration and failure over time, failure probability, reliability curves, risk of failure	Model description, background, input and output data, discussion	It is recommended to develop a predictive analysis to address gaps in the existing body of knowledge, focusing on the hydraulic conditions of the system and explicitly examining water velocity and pressure fluctuations

Table 2 (continued)

Source	Types of models considered	Type of prediction/nature of the output	Described items for each model	Final recommendations
Dawood et al. (2020a)	The review focused on Artificial intelligence methods: Failure prediction models, risk analysis models, water quality prediction models, condition monitoring and assessment models, remaining useful life models, leak detection and prediction models, and decision support system models	Annual breakage rates, time to next failure, failure rate, structural pipe capacity, pipe failure risk, probability and Consequence of failure, risk index of water mains, chlorine residuals, free ammonia levels, performance evaluation, corrosivity potential, remaining useful life	Model methodology, inputs, explanatory variables, algorithms, implementation examples, brief analysis, and outcomes	Models derived from artificial intelligence and machine learning are gradually gaining adoption. Nonetheless, there are still gaps in their implementation
Barton et al. (2022a)	The review focused on Statistical methods: deterministic, probabilistic, and machine learning	The number of failures, failure rates, probability of failure, failure age probability, time-to-failure, and failure priority-rank	Model methodology, general formula, advantages and disadvantages, analysis results, limitations, and comments	The limitation of utilizing restricted pipe failure inventories significantly impacts the quality achieved by complex models. Individual models per material yield better performance than global models

factors (Monfared et al. 2021), while another focuses on the statistical dependence of such models on explanatory variables (Gómez-Martínez et al. 2017). Additionally, a comprehensive examination of general-level factors affecting pipe failure in drinking water networks is presented in a separate publication (Barton et al. 2019).

In their research, Monfared et al. (2021) identified nine key factors or explanatory variables related to water quality data, including pH, chlorine (including free residual chlorine, chlorine decay, and chlorine concentration), temperature, turbidity, hardness, colour, water age, alkalinity, and conductivity. The study revealed a notable gap in understanding the precise influence of these variables on failure prediction models, highlighting the need for further investigation and analysis in this area.

Gómez-Martínez et al. (2017) examined thirteen explanatory variables' impact on their models. These variables encompassed physical characteristics such as diameter, year of installation, pipe material, and environmental factors like terrain type, land use, and depth of installation. Additionally, the study included hydraulic-related variables, namely pressures, velocities, and transients, which they referred to as internal variables. The research findings indicated that incorporating various explanatory variables did not yield significant advantages. On the contrary, simplifying the models by reducing the number of variables enhanced their reliability and facilitated interpreting results from a service-oriented perspective, particularly for water utility applications.

Although Barton et al. (2019) did not directly focus on failure prediction models, their research extensively examines the impact of physical, hydraulic, and environmental factors on the likelihood of failures, explicitly about drinking water pipes. The study is particularly notable for its comprehensive analysis of how these factors influence pipes, considering variations in materials and their distinct mechanical and chemical properties, which react differently to these variables. Moreover, the research underscores the significance of obtaining precise and comprehensive operational data and pipeline asset inventories to enhance the development of more accurate predictive models for pipeline failures and performance.

Despite the apparent contradiction between the latter two publications regarding the data required to train a model of this nature effectively, they refer to two different situations. The first study emphasizes that many variables can lead to overfitting, a situation where the model appears to fit well with the current data but fails to validate with future datasets due to the inclusion of noise variables (Belkin et al. 2019). On the other hand, the second publication stresses the importance of having a substantial and precise volume of inventory data for use in the models. These perspectives do not conflict, as a balanced number of explanatory variables

should be selected based on the specific conditions of each system under analysis and the chosen model.

The significance of Barton's insightful research (Barton et al. 2022a) is apparent as it showcases the evolution of statistical modelling by categorizing these models into three distinct types: deterministic, probabilistic, and machine learning. This research emphasizes the necessity for additional research to ensure the appropriate consideration of variables and other crucial factors in the accurate selection and execution of these models.

The current literature lacks a comprehensive understanding of how explanatory variables influence water pipe failure modelling within (WIAM). While previous research efforts have identified and analysed explanatory variables to some extent, there is a limited exploration of their impact on DWPF modelling processes. This review is based on the identified knowledge gaps, which include the absence of clear definitions of explanatory variables tailored to local conditions and specific modelling needs in DWPF models. Additionally, researchers need criteria to define these variables, and their influence on model selection and performance remains to be seen. Addressing these gaps is crucial for guiding future research.

To bridge these disparities, this research investigates methodologies from various studies, uncovering potential biases and limitations related to choosing and utilizing explanatory variables according to their model local conditions. By meticulously examining the literature, this research seeks to enhance the understanding of how different researchers attribute importance to explanatory variables and their effects on model performance. By shedding light on these aspects, the review will contribute to improving water infrastructure management and the performance of pipe failure prediction models.

Motivation and aim of this review

There are significant motivations to explore this field beyond the identified gap. Firstly, a transparent interconnection exists between water, energy, and food, with growing demands for these resources over time. Water plays a crucial role in all human processes as a fundamental element, underscoring the necessity to develop sustainable strategies for its utilisation (Carmona-Moreno et al. 2021). Leaks and failures in distribution systems are closely linked to pipe deterioration. Apart from the direct economic costs of repairing pipe failures, information on global Water Utilities' energy consumption often needs to be updated or updated. Additionally, other factors, such as the impact of failures and leaks on water quality, pose challenges in quantification (Chen and Guikema 2020).

As reference data, it is estimated that non-revenue water (NRW) losses accounted for 9.1×10^9 m^3 of water volume and 3100 GWh of energy loss in WSDNs in the USA in 2018 (Chini and Stillwell 2018). A comparison with the estimated energy consumption of 5600 GWh in the water distribution systems in the USA for the year 2005 (Mostafavi et al. 2018) indicates an approximate 44% reduction in the energy consumption required for WSDN operations. This reduction demonstrates the evident efforts made by Water Utilities to improve energy efficiency, despite the increase in drinking water consumption due to population growth. Implementing better strategies for hydraulic sectorization, optimizing pumping systems, advancements in pipe failure and leakage management processes, and other exogenous factors like per capita water use reduction have contributed positively to this trend.

The NRW index is a standard criterion for evaluating water distribution system performance, particularly concerning water leakage management. This index represents the difference between the volume of water supplied to the distribution system and the volume billed to consumers (Alegre et al. 2016). When the distribution network undergoes higher maintenance levels and exhibits improved integrity, the rate of annual pipe failures decreases, consequently leading to a reduction in NRW (Güngör-Demirci et al. 2018; Ananda 2019). Roigé et al. (2020) introduced the concept of water and energy losses as critical environmental criteria, highlighting the interconnectedness of service pressure, water leakage, and subsequent energy wastage. Figure 1 provides a graphical representation of the global status of pipe failures in distribution systems, showing a direct proportionality between higher rates of pipe failures and more significant NRW indices across various countries (Almheiri et al. 2021).

The environmental ramifications of pipe failures encompass greenhouse gas emissions resulting from non-optimized energy consumption in various water distribution processes affected by these failures and the water losses associated with such occurrences (Nair et al. 2014). In-depth studies, like those conducted by Herstein et al. (2009), outlined the economic, environmental, and social consequences of pipe production, installation, repair/renovation, and ultimate disposal. Additionally, these studies explore the implications of pumping processes, network pressure management, and hydraulic optimization in WSDN. While the current research may not cover the complete life cycle of the distribution system's diverse components (Herstein and Filion 2011), it facilitates a priori assessment of potential excess emissions attributed to network pipe failures.

Roigé et al. (2020) introduced several essential concepts related to water infrastructure, including the organoleptic perception of water, the risks associated with potential events and interruptions in drinking water service, as well as disruptions in pedestrian and motorized traffic. These parameters have a measurable impact on the prioritization of water pipeline renewal.

Although many authors have recognized water leakage in WSDN to cause significant social impacts (Gupta and Kulat 2018), the social aspects of the effects resulting from pipe failures, leaks, and water losses in WSDN remain relatively understudied, with limited research conducted on this subject beyond the work of Roigé et al. (2020).

Mazumder et al. (2021) evaluated economic, operational, environmental, and social consequences arising from the failure of integrated water and road segments, taking into account factors such as financial aspects (rehabilitation/renovation costs of pipes), operational indicators (service performance, hydraulic efficiency, road closures, and potential asset damage), as well as environmental and social impacts (effects on critical infrastructure, traffic, and population density).

Lee and Kim (2020) provided a summary of studies linking water leakages and other distribution network characteristics to sustainable development practices for WSDNs, encompassing economic, social, and environmental considerations. However, understanding the social implications of water leakages in WSDNs requires further investigation and exploration in the literature.

These studies consistently provide compelling evidence regarding the economic impacts of water pipeline failures. A substantial portion of a Water Utility's assets comprises pipelines, prompting numerous investigations focused on assessing the effects of water pipe failures from various angles: traffic congestion (Cunningham et al. 2021), pipeline characteristics (Mazumder et al. 2021), investment-based leakage reduction measures (Ahopelto and Vahala 2020), rehabilitation or replacement of failed water pipelines (Kleiner et al. 2010; Rahman et al. 2014), pipe replacement periods (Park 2011), and average network pressure (AL-Washali et al. 2020), among others. All the methods employed to estimate the economic impacts are based on functions and models that enable the analysis of potential benefits under various parameters and assumptions related to the technical management of the network.

Pipe failure modelling is a fundamental aspect of asset management models in WSDNs (Ugarelli and Sægrov 2022). It involves classifying pipe sections based on their likelihood of failure, which informs decision-making processes regarding renewal policies, maintenance strategies, and the identification of network sectors requiring further investigation for leak detection (Barton et al. 2021).

As mentioned in the preceding section, within the context of explanatory variables associated with water pipe failure modelling, a comprehensive literature review is required to synthesize previous studies, identify research gaps, and propose new directions for the field. To date, no study has
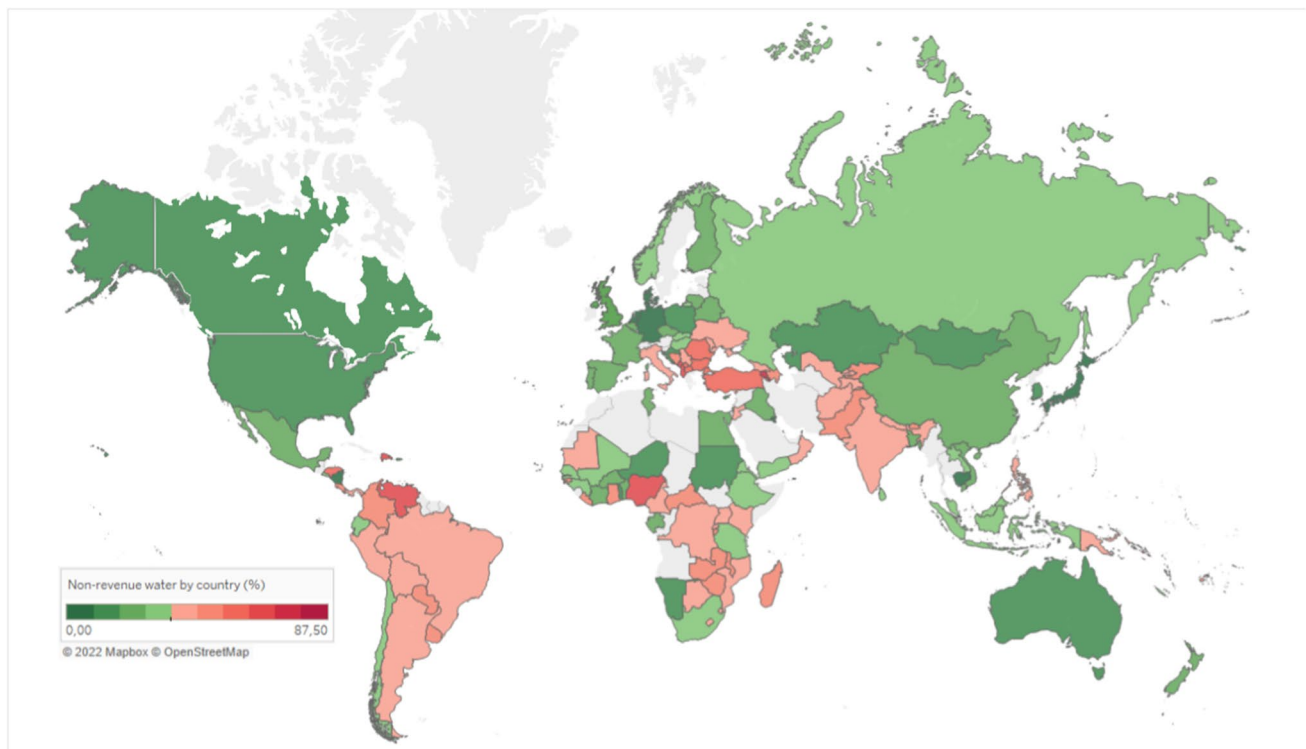

Fig. 1 Non-revenue water by country (%). Green tones indicate a lower NRW index; red tones indicate a higher NRW index. Data updated to the year 2021 (FP2E/BIPE 2019; OECD 2020; AEAS 2021; Go Associados 2021; Istituto Nazionale di Statistica 2021; Water and Sanitation Program 2021)

summarized pipe failure prediction models and their relationship with explanatory variables, underscoring the necessity to gain insights into the diverse perceptions of explanatory variables that may impact these models.

Outline of the work

This review aims to comprehensively present the various models employed in water pipe failure prediction, focusing on the explanatory variables utilized in models developed over the past 15 years. The primary novelty of this review lies in its exploration of how researchers attribute significance to different explanatory variables within their failure models. By examining the relationships between variables and the characteristics of the models, potential biases in the selection or utilization of explanatory variables can be identified. Moreover, we propose the inclusion of pertinent variables essential for the accurate execution of failure models in drinking water pipelines.

This paper is structured into five primary sections. The introduction provides an overview of Asset Management, linking it to water infrastructure, and emphasizes its relevance to water utilities. It underscores the significance of modelling water pipe failures to enhance the efficiency of water supply systems' operation and planning. Moreover, the introduction identifies a knowledge gap in these models, highlighting the necessity to improve the design and analysis of their explanatory variables.

"Review methodology and research questions" section outlines the review methodology and research questions, encompassing search terms, databases used, and exclusion criteria. "Models for prediction of drinking water pipe failure" section introduces the main categories of DWPF models under study. "Explanatory variables used for modelling drinking water pipe failures" section examines the critical explanatory variables that support the DWPF modelling, specifying the type of explanatory variable used, the employed model type, the model's output, and the time horizon of the model outcomes per publication. "Conclusions and future research directions" section concisely synthesizes the findings, providing valuable contributions to the topic, and outlines potential avenues for future research before concluding the review.

Review methodology and research questions

Review methodology

This review adopts a mixed approach, incorporating the postulates of Kitchenham and Charters (2007), Snyder (2019) and the PRISMA Guidelines (Moher et al. 2009). The review includes conference proceedings, journal articles, government documents, doctoral theses, and dissertations from 2007 to 2023, sourced from five academic databases: JSTOR, EBSCO, ProQuest, Scopus, and Web of Science. Initially, a search with the phrase "Drinking water pipe failure modelling" yielded 2914 results, further refined using specific search strings and assessed against the exclusion and inclusion criteria outlined in Table 3. Following this process, the research narrowed to 103 relevant manuscripts, as depicted in Fig. 2.

The definition of failure, as employed by Le Gat (2015), plays a crucial role in delineating the scope of this research. The water infrastructure under consideration is structured as a network of interconnected pipelines and failures, encompassing leakage or breakage, typically manifest clustered, affecting specific network segments.

Research questions

This paper aims to conduct a meticulous literature review on water pipe failure models, primarily focusing on the foundational explanatory variables underpinning their analyses and outcomes. The study's objectives are framed by eight research questions (RQ), as outlined in Table 4. These questions define the study's scope and provide a roadmap for gathering and analysing relevant information, enabling the exploration of challenges and potential advancements in this domain. The Introduction section addresses the first three questions, setting the main trajectory of the review and justifying its necessity while establishing its link to asset management.

Models for prediction of drinking water pipe failure

A brief overview of modelling for the prediction of drinking water pipe failure (DWPF)

Employing a water pipe failure prediction model enables estimating future break/failure events based on historical observations, determining an appropriate renewal rate, and supporting decision-making processes related to key indicators, renewal scenarios, and selecting pipes for replacement. In recent decades, water pipe failure modelling has become a valuable tool for analysing failure data collected from WSDNs, serving as a standard planning approach to investigate potential causes of pipeline failures. Initially, Water Utilities relied on expert judgement to characterize failure events, considering factors such as pipe age and applied pressure. However, the evolution of this approach has aimed to enhance the information captured regarding

network failures and improve analysis through various statistical techniques.

The progress in failure event studies has facilitated the identification of specific scenarios and requirements, leading to more effective proposals for network renovation or reinforcement. By improving the quantity and quality of collected information, new indicators can be generated, and existing ones refined, thus impacting failures. Water Utilities are increasingly focused on predicting pipe failures or deficiencies before they happen, evident in the analysis of failure rates per pipe. This involves considering factors related to existing defects, such as previous failures, leakages, and ageing, as well as potential improvements, to prioritize and select pipes for renewal based on their probability of failure. However, a significant challenge in implementing this approach lies in the availability of comprehensive historical data regarding WSDNs, which includes factors such as pipe material, location, age, and failure history.

The emergence of machine learning techniques has significantly enhanced statistical models focused on studying pipe failure phenomena. This development proves beneficial in addressing the limited availability of historical data faced by Water Utilities when investigating trends and probabilities associated with pipe failures. Using sophisticated algorithms, machine learning techniques can effectively estimate missing data by leveraging the stochastic nature of the missing values within the WSDN information dataset. Nonetheless, applying these techniques demands a high level of reliability in the historical information of the WSDN to ensure it adequately represents the entire data universe.
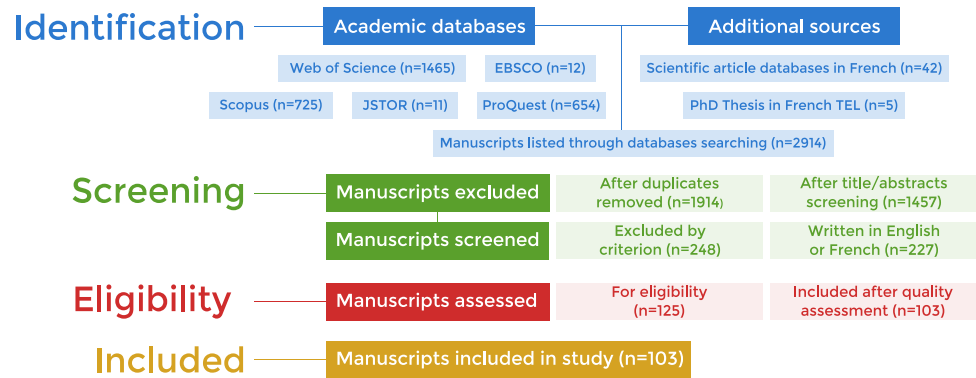
While this paper exclusively focuses on water pipe failure models for WSDN and their associated explanatory variables, it is crucial to acknowledge the progress made in pipe-related research for other purposes. This includes investigations into the mechanical or rheological properties of piping materials. Although such studies fall outside the scope of this paper, it is worth noting that the models presented herein draw upon the knowledge generated in these areas to enhance the predictive capabilities of failure models within the context of multi-criteria analysis conducted by each analysed model.

Classifications of models applied to the investigation of pipe failures in drinking water systems

In recent decades, considerable interest has been in predicting DWPF as WSDN gradually deteriorate. Building upon the methodologies employed in previous reviews (Dawood et al. 2020b; Karimian et al. 2021), this study updates the list of recent research on DWPF models in Table 5 The table systematically categorizes and provides a concise summary of all the models employed by researchers in the last fifteen

Table 3 Inclusion and exclusion criteria for including manuscripts in this review

ID	Inclusion criterion	Exclusion criterion	Observation
1	The manuscript must have been published between 2007 and 2023	Publications before 2007 or after March 1, 2023	The preceding research conducted before 2007 has generated valuable knowledge that subsequent publications have incorporated
2	The manuscript must focus on failure modelling for drinking water pipes/networks	Any other type of pipeline (e.g. oil pipelines, wastewater, nuclear applications)	Restriction due to the relevant economic weight of other pipeline operations, which generates research in these fields
3	The manuscript must focus on modelling studies using data from a drinking water network	Models, from a purely mechanical or physical perspective, of rheological phenomena and materials' strength	A crucial aspect of the excluded publications involves studying explanatory variables and their influence on pipelines through physical testing conducted in laboratory environments
4	The manuscript must be written in English or French, peer-reviewed and through a recognized publisher	Non-peer-reviewed studies presented in other languages or with no possibility of external validation	French has been included as an accepted language of dissemination due to its significant contribution from French-speaking researchers in this field

Fig. 2 The review process, PRISMA flow diagram

Table 4 Research questions

ID	Research question	Description
1	What is the current application of asset management in WSDN?	The importance of water infrastructure in society's well-being underscores the need for effective Asset Management, making it crucial to investigate their relationship. The initial two research questions are addressed in the Introduction
2	Why are models crucial in WIAM, and what types of output do they provide, at different levels?	Models are indispensable in WIAM, and their outputs exhibit significant heterogeneity, warranting thorough investigation and explanation
3	What are the current gaps in the literature concerning the linkage to explanatory variables, and what specific requirements does this review aim to fulfil?	While there are existing studies on failure prediction models, there needs to be more synthesis and analysis that focuses explicitly on the role of explanatory variables in these models
4	What models predict WSDN pipe failures?	This research aims to comprehensively examine models used in predicting water pipe failures within Water Utilities
5	What are the evolving trends in the field of DWPF modelling?	By examining model usage trends and their associated explanatory variables over time, this study aims to analyse the underlying reasons driving these developments
6	What explanatory variables serve as the foundational information for modelling DWPF in WSDN?	This study seeks to compile and present the frequency of usage of explanatory variables that support the models under consideration
7	What criteria influence the decision to incorporate explanatory variables into the processes?	This study aims to enhance researchers' understanding of the critical role played by identifying and defining explanatory variables in the modelling process, with specific consideration of the unique conditions present in each Water Utility
8	What potential future research directions exist considering DWPF modelling?	This research seeks to offer valuable insights and improvement opportunities to researchers, developers, and Water Utility professionals involved in developing accurate failure prediction models for drinking water pipes

years, directly addressing research question (RQ$_4$) (Table 6). The information presented in this table has been extracted from all the sources listed in Table 7.

The main classifications are derived from the significant differences observed in the approaches commonly used in the analysed research. Statistical-based models seek to establish relationships between variables through mathematical equations (Fahrmeir and Tutz 2001). In contrast, Machine Learning-based models take a different approach, utilizing algorithms to learn from the data and establish these relationships (ICAMLDA 2010). Other proposed classifications

deviate from these definitions, exploring alternative paths for failure estimation.

Comments on the limitations of current models

The necessity of utilizing historical pipeline inventories and failure data gives rise to a well-recognized issue of left-truncation in executing statistical-based survival models, owing to the nature of event-related data over time. This phenomenon, along with right-censoring, has been extensively examined by Le Gat (2015) and other authors

(Robles-Velasco et al. 2021; Barton et al. 2022b), underscoring the significance of mitigating potential interferences in predictions and time estimations to determine the first failure (Xu and Sinha 2021). As aptly pointed out by Scheidegger et al. (2015), left-truncation and survival selection are prevalent features in the available data concerning urban water services. Nonetheless, a comprehensive analysis of the impact of these features on modelling processes is lacking in most reviewed studies, and their direct consideration remains limited to date.

Machine learning models offer several advantages over statistical models, including their improved response to outliers and capacity to establish meaningful relationships between explanatory variables and pipe failures, thus defining the significance of each variable in the modelling process. Despite the growing popularity of machine learning-based models, it is crucial to acknowledge that they may need help incorporating right-censored information, potentially leading to overestimating pipeline faults beyond their actual occurrences (Snider and McBean 2021).

There are other disadvantages associated with machine-learning-based methods. One of the main drawbacks is that these methods, by nature, lack physical constraints. Unless they are explicitly imposed with specific conditions, they do not inherently consider the limitations of the physical environment. As a result, it becomes necessary to critically assess the results of such models, given the challenge of internally verifying their implementation.

Another intricate issue with such models is their interpretability (Barton et al. 2022a). These models can become effective with interpretability, as human interpretation involves considerations beyond the technical proficiency of the modelling process. Commonly used techniques, like SHAP or LIME, are employed in other fields to explain predictions made by such models; nevertheless, in this field, few studies utilize these methods (Fan et al. 2021). The outcomes of these analyses sometimes align with expected or observed results in the field, and they rely on an in-depth understanding of the variables' components and effects specific to each environment and network.

Using standard metrics such as RMSE and MCC may not inherently reflect the practical value of projections made by a machine-learning-based model concerning service needs. An example of this occurs when a model achieves high standard metrics, but its performance may need to be more optimal when the Lorenz Curve analysis is applied (Le Gat 2015), which significantly illustrates the impact of the models on renewal needs.

Among the advantages and disadvantages outlined by Barton et al. (2022a) and Almheiri et al. (2020a), the necessity for hyperparameter tuning in certain cases renders the implementation of such models highly demanding.

Furthermore, it is essential to consider the computational power requirements highlighted by Gupta and Segal (2022), as the choice of model type depends on both the service needs, the scope of the DWPF modelling projects, and the utility of their outputs for stakeholders.

In addition to variable types, time and spatial frameworks, the type of response and level of inference, as detailed by Barton et al. (2022a), the focus of the model application may also influence the models' implementation conditions. For instance, a model aimed at long-term planning would include distinct modelling capabilities, with survival analysis models being commonly selected in such cases.

The nature of data related to pipeline failure events is inherently unbalanced, as in most cases, only 0.1% of the data universe contains one or more failures (Barton et al. 2022b). This significant imbalance underscores the requirement for comprehensive records of failures over an extended period to ensure that the limited data can offer sufficient representativeness for unbiased analysis. An insufficient number of periods considered may lead to underfitting in machine-learning-based models, where the model needs more information due to either high bias or excessive variance. This limitation is also evident when there is an inadequate amount of data for each feature in the analysis, as it is essential to have sufficient training instances to adjust the models effectively.

A prospective area for future research entails addressing the limitations related to data acquisition within Water Utilities, aiming to ascertain the appropriate proportions and data volumes required to implement a failure model effectively. Investigating this direction would encompass various factors, such as the network's characteristics, the availability of inventories or changes within the system, and the users' specific requirements.

Data imbalance constitutes one of the most significant challenges faced by various service operators, as collecting this information was only integrated into Water Utilities' procedures relatively recently. Addressing the issue of data imbalance can be achieved through the utilization of synthetic samples (Robles-Velasco et al. 2023). This approach proves beneficial when faced with limited information, thus mitigating the challenges posed by significant imbalances in the data. Caution must be exercised to avoid generating excessive synthetic samples, which may lead to suboptimal model performance.

The optimization of calibration thresholds plays a crucial role in mitigating the impact of utilizing imbalanced data by striking a balance between sensitivity and specificity (Esposito et al. 2021; Barton et al. 2022b). Additionally, sampling methods, such as stratified sampling, can address this imbalance and ensure adequate representation of materials in both the training and test datasets (Winkler et al. 2018; Barton et al. 2022a).

Table 5 Models applied to analyse and predict pipe failures in drinking water systems, compilation to date

Primary classification	Secondary classification	Technical foundation of the modelling method—Selected source	Model	WSDN length under study (km)
Statistical-based models	Stochastic Modelling and Survival Analysis	Proportional hazards models—Bayesian Weibull -WPHM (Snider and McBean 2021)	M1	4387
		Proportional hazards models—Cox -CHM- (Almheiri et al. 2021)	M2	1195
		Bayesian Belief Network—BBN (Giraldo-González and Rodríguez 2020)	M3	1819
		Weibull distribution—WD- (Weeraddana et al. 2020)	M4	1888
		Linear Extension of the Yule Process with selective survival—LEYP2s (Le Gat 2015)	M5	721
		Linear Extension of the Yule Process—LEYP- (Renaud et al. 2012)	M6	550
		Non-Parametric Estimation/Kaplan–Meier estimator (Le Gat 2015)	M7	721
		Vector autoregression—VAR- and vector autoregression with exogenous variables -VARX- (Almheiri et al. 2020a)	M8	–
		Hierarchical Beta Process—HBP (Chik et al. 2017)	M9	376
		Markov model/Markov Chain Monte Carlo—MCMC- (Osman and Bainbridge 2011)	M10	1593
		Monte-Carlo Simulations—MCS (Davis et al. 2007)	M11	–
		Hybrid estimate using Herz distribution (Le Gat et al. 2013)	M12	–
	Regression analysis	Multiple Regression Model—MRM (Wang et al. 2009)	M13	432
		Poisson Regression Model -PRM (Asnaashari et al. 2009)	M14	56
		Non-homogeneous Poisson process -NHPP- (Chen and Guikema 2020)	M15	681
		Gaussian Process Regression -GPR- (Weeraddana et al. 2020)	M16	1888
		Generalized linear/non-Linear model -GLM-, -GNLM- (Chen and Guikema 2020)	M17	681
		Generalized additive model -GAM- (Chen and Guikema 2020)	M18	681
		Linear mixed-effects model -LMEM- (Chen and Guikema 2020)	M19	681
		Multivariate Adaptive Regression Splines -MARS- (Aslani et al. 2021)	M20	3476
		Time Linear Model -TLM- (Konstantinou and Stoianov 2020)	M21	374
		Time Exponential Model -TEM- (Konstantinou and Stoianov 2020)	M22	374
		Multi-objective Genetic Algorithm and Evolutionary Polynomial Regression -EPR- (Karimian et al. 2021)	M23	5045
		Bayesian linear regression -BLR- (Gómez-Martínez et al. 2017)	M24	17,473
		Ridge Regression -L2 Regularization Method- (Almheiri et al. 2020b)	M25	432
	Probabilistic	Survival analysis – Parametric (Christodoulou et al. 2010)	M26	795
		Survival analysis – Nonparametric (Christodoulou et al. 2010)	M27	795
		Naïve Bayes algorithm -NB- (Konstantinou and Stoianov 2020)	M28	374
		Bayesian network/Bayesian simple model -BSM- (Tang et al. 2019)	M29	342,850
		Linear discriminant analysis -LDA- (Konstantinou and Stoianov 2020)	M30	374

Table 5 (continued)

Primary classification	Secondary classification	Technical foundation of the modelling method—Selected source	Model	WSDN length under study (km)
ML- and AI-based models	Artificial neural networks	Advanced Meta-Learning -AdvaML- (Almheiri et al. 2021)	M31	1195
		Multilayer Back-propagation Neural Network -BPNN-/Multilayer Perceptron -MLP- (Kerwin et al. 2020)	M32	411
		Extreme Learning Machine -ELM- (Sattar et al. 2019)	M33	971
		General Feed Forward Neural Network -FFNN- (Aslani et al. 2021)	M34	3476
		Radial-Based Function Neural Network -RBFNN- (Jafari et al. 2021)	M35	80
		Generalized Regression Neural Network -GRNN- (Aydogdu and Firat 2015)	M36	440
	Supervised learning	Least Squares Support Vector Machine -LS-SVM- (Aydogdu and Firat 2015)	M37	440
		Support Vector Machine/Regression/Classification -SVM- (Almheiri et al. 2021)	M38	1195
		Random Forest -RF- (Snider and McBean 2021)	M39	4387
		Survival Random Forest -SRF- (Almheiri et al. 2021)	M40	1195
		Boosted Decision/Regression Trees -BRT- (Aslani et al. 2021)	M41	3476
		Gradient-Boosted Trees, Networks and Frameworks -Gradient boosting- (Chen and Guikema 2020)	M42	681
		Ensembles, Decision Tree -EDT-, Bagging, Boosting (Almheiri et al. 2020b)	M43	432
		Bayesian Model Averaging -BMA- (Demissie et al. 2019)	M44	1480
	Clustering	Classification: K-means/K-nearest neighbours -KNN- (Giraldo-González and Rodríguez 2020)	M45	1819
		Cluster-weighted modelling -CWM- (Chen and Guikema 2020)	M46	681
Fuzzy logic-based models	Fuzzy system	Hierarchical Fuzzy inference system -HFIS- (Ward et al. 2017)	M47	–
		Adaptive Neuro-Fuzzy Inference Systems -ANFIS- (Tabesh et al. 2009)	M48	579
		Mamdani method (Fares and Zayed 2010)	M49	153
Lifecycle analysis models	Lifecycle cost	Life Cycle Cost Analysis -LCCA- (Francisque et al. 2017)	M50	216
Evolutionary computation models	Evolutionary algorithm	Gene Expression Programming -GEP- (Sattar et al. 2016)	M51	1021
Geospatial Data Analytics models	Satellite observations	Pixel-based approach—Soil deformation analysis -SDA- (Arsénio et al. 2015)	M52	4309
Decision-making models	Multiple criteria	Analytic Hierarchy Process -AHP- (Al-Barqawi and Zayed 2008)	M53	–
Risk analysis models	Probability-consequence model	Weighted risk analysis -WRA- (Barton et al. 2022b)	M54	38,424

Despite the assertion that machine-learning-based models do not effectively address the issue of data imbalance, leading to lower accuracy in failure prediction (Robles-Velasco et al. 2021), it is worth exploring the solutions

---

Table 6 Key features compendium of relevant recent research in pipe failure prediction models

References	Key notes	Primary aim of the research	Noteworthy insights
Almheiri et al. (2021)	A novel ML model is proposed, applying the meta-learning paradigm, to predict the Likelihood of Failure (LoF). The model considers a comprehensive set of variables, surpassing conventional studies in this domain	This research seeks to introduce a novel LoF prediction model suitable for application in case studies with partial, incomplete, or insufficient data. The model aims to predict the hazard ratio of pipe "i" that has not experienced a failure event during the condition monitoring period of water distribution pipelines at time "t."	In the context of this study, residual chlorine and the number of traffic lanes emerged as the most influential factors impacting the LoF. However, it is noteworthy that the research does not specify the time range covered by its predictive outcomes, which focus on the service life length of water pipes
Karimian et al. (2021)	This research used the EPR method to predict the number of water pipe failures by incorporating multiple explanatory variables	In addition to conducting a comprehensive literature review, this research endeavours to develop a predictive model for water pipe breaks and perform a sensitivity analysis on this model	This research conducted a sensitivity analysis of the explanatory variables to gain insights into the anticipated changes in the predicted number of pipe failures. The study investigated the interrelationship between material variations, standards, and the outcomes of the sensitivity analysis
Snider and McBean (2021)	Pioneering integration of survival analysis techniques within a Machine Learning framework, effectively addressing the limitations identified in prior research	A precise long-term pipe replacement model was developed by integrating right-censored break data into its calibration	The results indicate that the RSF model outperforms traditional ML and survival analysis models in predicting pipe failures, particularly for long-term pipe replacement strategies, specifically for ductile iron pipe typology
Aslani et al. (2021)	Spatial clustering was employed to identify areas vulnerable to pipe failures. The results were then used to compare the failure rates using four machine learning algorithms—RF, BRT, MARS, and ANN—for the WDN in Tampa, USA	This research aimed to develop predictive models for water main failures using spatiotemporal data while incorporating additional variables into the standard database, such as average temperature and total precipitation	One of the paper's notable findings is that the environmental conditions included in the models, namely average temperature and total precipitation, do not significantly impact the variables to be considered in the analysed models
Barton et al. (2021)	Provides an in-depth examination of water professionals' and experts' perspectives through interviews, focusing on the limitations and opportunities of pipe failure models	This study aims to enhance the understanding of critical challenges, uncertainties, and potential improvements for pipe failure models across five categories: data management, data limitations, pre-processing difficulties, model scalability, and future opportunities	Using low-quality or insufficient data sets, coupled with the imbalance between model-projected and actual failures, results in inaccurate short-term failure modelling outputs for individual assets. Consequently, such inaccuracies can lead to erroneous decisions in the planning process
Chen and Guikema (2020)	Spatial clustering of pipe failures using three algorithms: Poisson-based, density-based, and locally weighted density-based. Investigation and validation of pipe failure data clustering in a real-world water distribution network	This study aims to demonstrate that high-precision clusters can offer valuable insights for more accurate predictions of future pipeline failures. The research findings indicate that spatial groups formed based on historical break events can contribute significantly to predicting future pipe failures	The conclusions drawn in their study are derived solely from empirical outcomes and are specific to their dataset. Consequently, the applicability of employing spatial clusters may differ in various test cases

Table 6 (continued)

References	Key notes	Primary aim of the research	Noteworthy insights
Weeraddana et al. (2020)	Introduces a semi-parametric Bayesian -Gaussian Process Regression- model for pipeline failure. Forecasting, capturing the survival probability and handling covariates	In forecasting water pipe failures, models that attempt to learn from a limited number of data points may suffer from overfitting issues. The proposed approach in this study addresses the limitations observed in commonly employed Machine Learning methods	Based on the tests conducted in several Australian Water Utilities, this model demonstrated superior performance compared to other Machine Learning-based models
Snider and McBean (2020)	Compares two modelling methods: A gradient-boosting decision tree ML model and a Weibull proportional hazard survival-analysis model. Forecasting pipe breaks in a Canadian WDN	The primary objective of this study is to describe and apply an XGBoost model and a Weibull proportional hazard model for comparing their time predictions to the next failure event for individual cast-iron pipes within the WDN. Additionally, the research aims to investigate the effect of including censored data in the models	The findings indicate that Machine Learning (ML) models are more suitable for short-term pipe-break modelling. However, it is essential to note that excluding censored events during ML model development can impact the accuracy of long-term predictions
Dawood et al. (2020a)	A comprehensive literature review was implemented to generate a state-of-the-art AI-based deterioration modelling for urban water systems	The primary goal of this research is to review pipeline failure modelling and evaluation alongside artificial intelligence and soft computing applications, with suggestions for future research	Application of ML and AI-based models for water system management is scarce, as the predominant research centred on physical and statistical modelling approaches and sporadic AI-based inclusions
Giraldo-González and Rodríguez (2020)	This study compares four ML-based models -GBT, Bayes, SVM and ANNs- with three statistical-based models—linear, Poisson and EPRs—in predicting individual pipe failure rates -PFR- for a case study in Colombia	The performance of the models above was evaluated based on pipe diameter, age, and length, with some environmental and operational variables. The Poisson and GBT approaches were the most suitable for predicting failures for the test sample of asbestos-cement and PVC pipes	The significance of developing pipe failure models for every independent network is highlighted, considering that the differences and characteristics of each network may impact specific variables to be considered by the models
Konstantinou and Stoianov (2020)	This paper proposes a comparative study of statistical and ML techniques to conduct an exploratory causal analysis of historic pipe breaks in a UK operational water supply network	Investigate how various physical phenomena affect the probability of individual and cohort pipe breaks, also classifying pipes based on their vulnerability to failure	Although evaluated models present insights into the causes of historic pipe failures, their analysis also highlighted that data quantity and quality must be thorough -such as water temperature and information about the failure properties-. The research does not describe the time range supported by its predictions
Kerwin et al. (2020)	This research includes pipe material as an input parameter for an ANN model to estimate the TtNFE, addressing right-censored data by combining generated failures data with observed failures data	This paper examines the impact of including failure data generated to fill in the missing information in ANN's failure prediction model for short-term prioritisation	The authors claim that the pipe break prediction model's overall performance is not affected by integrating failures generated for newer materials
Tang et al. (2019)	Bayesian network models for estimating the LoF in WDN were assembled by automated learning and a manual process based on a literature review and exploratory data analysis	This research aimed to compare two methods to develop Bayesian networks -an automated data-driven method and a guided method utilizing prior knowledge and conventional data exploration- to model the LoF for various pipe materials	Even with the automated method through ML obtaining slightly better accuracy than the manual model, its output presented a complex interpretation structure

Table 6 (continued)

References	Key notes	Primary aim of the research	Noteworthy insights
Shirzad and Safari (2019)	This paper compares two ML techniques –MARS and RF– to predict pipe failures for two WDNs, considering pipe information –diameter, length, depth, age, and average hydraulic pressure– as input variables	The objective is to compare the accuracy of both ML models to predict the PFR in two real WDNs and identify the most important variables affecting the PFRs for these case studies	In the two study cases, the variable of pipe installation depth was discarded for the two methods analysed as it was not identified as a variable that affects the results of the models. The research does not describe the time range supported by its predictions
Chen et al. (2019)	The research examined whether applying statistical models without using data from the WDN pipes, but considering other variables and the failure history, can still provide helpful information to guide asset management despite its limited accuracy	One of the study's objectives was to determine whether a classification approach based on the predictions of a statistical model can reach greater accuracy than one based on historical failure rates for this specific case	In compliance with the authors, despite the limited accuracy of models that do not account for detailed pipe data, they may better prioritize high-risk assets than a historical ranking based on failure rate. The research does not describe the time range supported by its predictions
Sattar et al. (2019)	This research presents an ML approach –extreme learning machine– to predict the TtNFE of individual pipes for a WDN	This research compares the developed model's precision and another algorithm based on ML –feed-forward ANN trained by backpropagation, support vector regression, and nonlinear regression–	The proposed model includes the pipe protection method as a variable and shows how the pipe coating with cement mortar and cathodic protection could significantly increase pipe lifespan. The research does not describe the time range supported by its predictions

proposed by Kaur et al. (2020). Chen et al. (2019) suggested increasing the spatial resolution of the data, resulting in a loss of accuracy in predicting non-failure events but an improvement in failure predictions.

An essential aspect of model generation lies in the necessity for expert knowledge to define model parameters and conduct data preparation processes. This expertise allows for assessing the models' predictions concerning real-world field conditions. Moreover, expert knowledge facilitates the inclusion of external factors that may be challenging to quantify and integrate into the models, leading to manual parameter adjustments (Barton et al. 2022b) to align the results with the specific context of the Water Utility concerning budget and strategy.

Given the diverse characteristics of each WSDN inventory, direct comparisons of model performance metrics across different networks become problematic because these metrics heavily depend on the quality and size of the inventory on which the models are based (Robles-Velasco et al. 2021). While these inventory differences hinder the ability to compare metrics directly, it is possible to compare different network models under specific conditions where data availability limitations can be overcome.

In cases where complete and reliable inventories are available as the base data for the models, it becomes feasible to compare different network models. To achieve this, algorithms purely based on this comprehensive data are utilized. Additionally, in the case of machine-learning-based models, comparisons are possible when the same algorithm, configured with the same parameters and hyperparameters, is applied. Through such rigorous comparative approaches, the metrics of these models can be effectively compared.

Expert criteria are paramount in model development, particularly in aiding the selection of the appropriate model type to align with the specific needs of the WSDN system. Understanding the stakeholders' requirements necessitates collaboration between system operators and model developers, mainly when they belong to different domains. For machine-learning-based models, selecting hyperparameters presents a complex task, and their variability can significantly impact the model's outcomes. This selection process is computationally expensive, as the effects of hyperparameters become evident only after executing all model processes. Therefore, validation and feedback from system experts are crucial.

While specific tools exist to optimize hyperparameters and enhance model performance automatically, they have yet to deliver optimal results in environments with highly unbalanced data (Czako et al. 2021), which is characteristic of our particular case study.

By defining the weights of each criterion, these experts can tailor the model to the specific requirements of each WSDN (Assad and Bouferguene 2022). The absence of

expert judgement may include irrelevant explanatory variables that do not influence pipeline failure processes, leading to suboptimal selection of a reference model or misinterpretation of model outcomes. Such misinterpretations can lead to misguided investments directed towards infrastructure within their amortisation period (Almheiri et al. 2021).

Robles-Velasco et al. (2021) assert that machine-learning-based models with more precise pipeline failure predictions tend to exhibit reduced non-failure forecasts. Barton et al. (2022b) explain how this challenge can be addressed through appropriate threshold optimization—a critical decision-making step unique to each Water Utility, necessitating expert judgement.

An inherent challenge in model generation and optimization is the selection of suitable explanatory variables. One approach to tackle this is by analysing small groups of variables to assess their significance (Robles-Velasco et al. 2021). In the case of medium-sized WSDNs, reducing the number of explanatory variables helps mitigate the risk of model overfitting (Jenkins et al. 2015).

Establishing the number of explanatory variables is a priority process (Fan et al. 2014), and it can be distilled into three constraints: the relevance and redundancy of the variable, the availability of computational resources and time for model execution, and the interpretability of the model outcomes enabling the understanding of complex relationships between the variables involved. The evident necessity for a comprehensive study encompassing the explanatory variables used in the models further reinforces the rationale behind this research.

Table 6 summarizes critical features from selected studies on the subject. It outlines their outputs' scope, objectives, and nature and presents meaningful insights specific to each case study.

Evolving trends in modelling for predicting DWPFs

The subsequent three figures and their corresponding concepts are based on developments derived from the open-source "bibliometrix" R-package (Aria and Cuccurullo 2017). Figure 3 depicts the conceptual structure of the research topic, highlighting the interconnections between the concepts used in the titles and abstracts of the relevant manuscripts. The size of the circle and the text in the figure represent the current importance of each concept, as determined by applying the Fruchterman–Reingold algorithm (Aria and Cuccurullo 2017).

The figure reveals a lack of uniformity in the use of terms such as "water main," "pipe," and "pipelines," as well as "breaks" and "failure". The analysis suggests that five key concepts unify the research into distinct clusters, namely the analysis of pipeline failures and

infrastructure asset management through machine learning and modelling.

Figure 4 highlights the most frequently used keywords in the studied manuscripts over the past 15 years, illustrating their evolution. Notably, the term "corrosion," which was once considered a crucial explanatory variable, has seen a decline in usage, while "replacement" and "patterns" have also diminished in popularity. Conversely, terms associated with applying artificial intelligence and machine learning methods are gaining prominence.

The significance of critical concepts is depicted in Fig. 5 using the Sankey diagram (Aria and Cuccurullo 2017). The diagram illustrates these concepts' evolution, with the rectangles' size representing their relevance during the specified period. Notably, the "Statistical analysis" concept has not seen new developments in the last two years. On the other hand, "Asset management" and "Pipe Failure" concepts have paved the way for "Machine learning" and "Data Mining." Additionally, the diagram indicates that Bayesian model averaging has yet to experience recent advancements. This chapter confirms the paramount importance of machine learning techniques in the research topic, effectively addressing RQ_5.
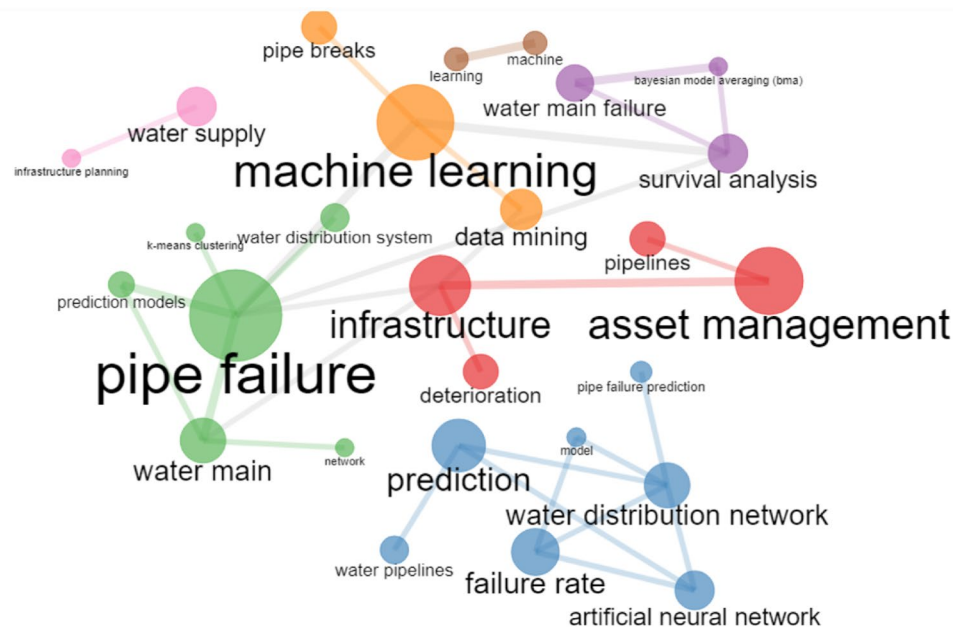
The term "Data mining" does not signify a new trend but rather its repeated mention in articles involving machine learning techniques. The interpretation of Fig. 5 reveals that the focus of applied research has shifted from using complex statistical models to harnessing sufficient computational power for implementing and testing various machine learning techniques. This trend is corroborated by the work of Barton et al. (Barton et al. 2022a), which outlines the evolution of DWPF models from statistical applications to the adoption of machine-learning-based models, as illustrated in Fig. 6.

This trend overlooks the significant potential that statistical models hold. As previously noted, statistical-based models directly incorporate external knowledge into the model, a crucial difference from machine learning-based models that require adjustment parameters governing the incorporation of superficial knowledge (Binder 2014). Such models need more extensive development, which might not be replicable in other systems without a similar adjustment phase.

It is essential to acknowledge that the concepts depicted in Fig. 5 do not solely dictate the current and future trends in model evolution. With advancements in modelling techniques, there is a greater capacity to analyse uncertainties and sensitivities associated with the employed variables.

Most studies that address modelling failure for drinking water pipes and uncertainties are predominantly based on Monte Carlo simulations (Beh et al. 2017; Jafari et al. 2021; Fan et al. 2023) and Poisson distributions (Xu et al. 2018). However, these methodologies have also been extensively employed in various related disciplines, such as hydraulic

Fig. 3 The conceptual structure of the modelling of DWPF for the last 15 years

modelling (Braun 2019), optimization in water infrastructure planning (Beh et al. 2017) and flood damage assessments (de Brito et al. 2019; Morita and Tung 2019) among numerous other applications, which illustrates the importance of analysing uncertainties in modelling processes.

These concepts continuously evolve and provide fertile ground for extensive research and exploration. Research predicting or forecasting events based on highly unbalanced data, such as in our case, emphasizes the utmost importance of data quality. Hence, mitigating the influence of uncertainty throughout all stages of the modelling process will not only enhance the reliability of the models but also minimize potential biases inherent in the analysis (Fan et al. 2023).
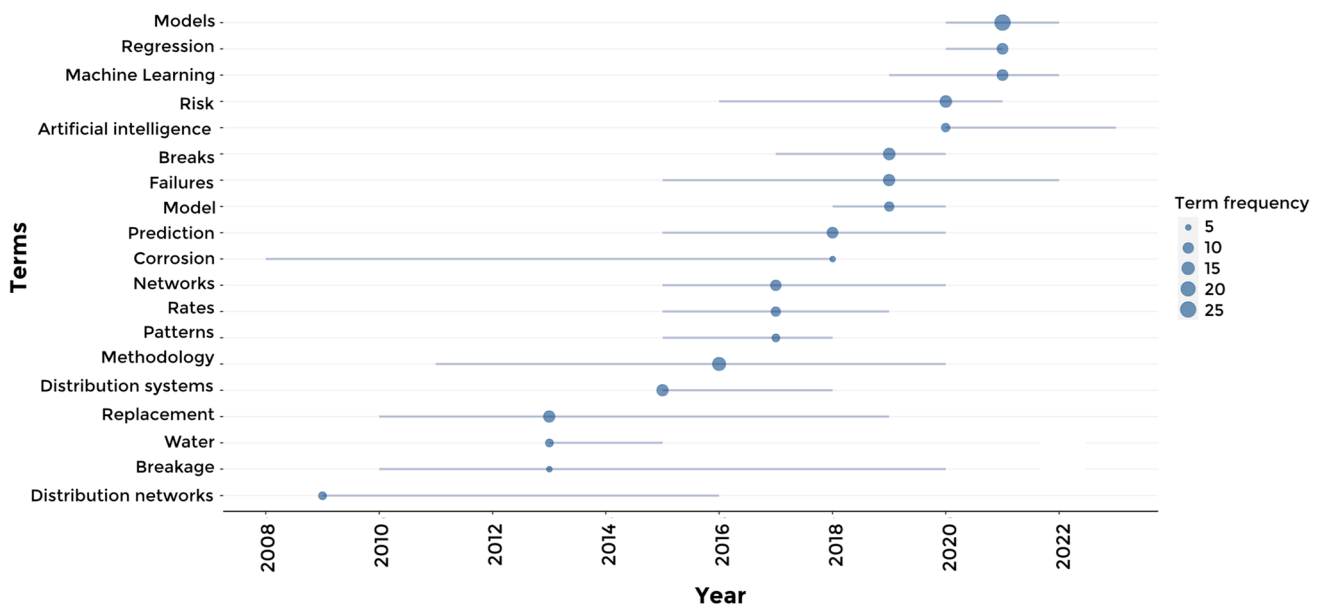

Fig. 4 The presence of research keywords from 2007 to 2023. The blue line's length represents the years in which the publications mention the keywords. The position of the blue ball indicates the year with the highest frequency of mentions for each keyword, and the size of the ball reflects the frequency of mentions in that particular year
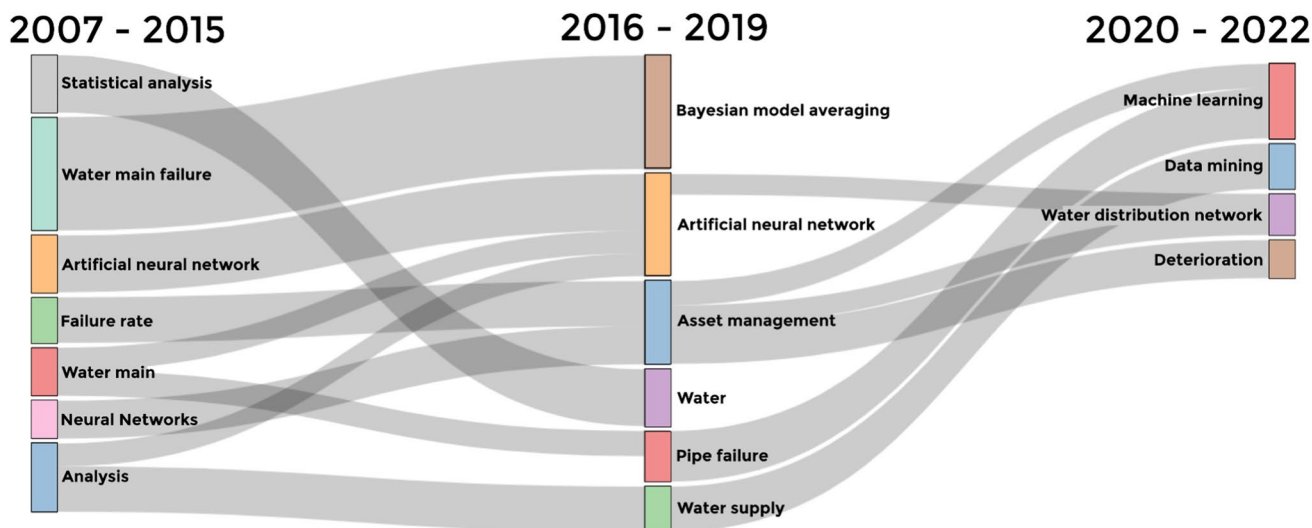
Fig. 5 Evolution of critical concepts researched for the periods 2007–2015 -left-, 2016–2019 -centre- and 2020 to 2022 -right-

Explanatory variables used for modelling drinking water pipe failures

Table 7 in Appendix 1 summarizes the explanatory variables extracted from the publications from 2007 to early 2023. This table establishes a link between the model typology proposed in Table 5 and the models implemented in each study, enabling an analysis of the selected model trends. Additionally, it outlines the optimal performance exhibited

by each model in the respective studied network. The table also indicates the type of output generated by each study, encompassing failure probabilities, estimated failure times, failure prediction focus, and the creation of performance or failure-related risk indexes and curves.

Table 7 also presents a detailed account of the primary explanatory variables employed in the modelling process. It delineates the type of explanatory variable utilized, the corresponding model output, and the time horizon for the model outcomes in each publication. Through this comprehensive

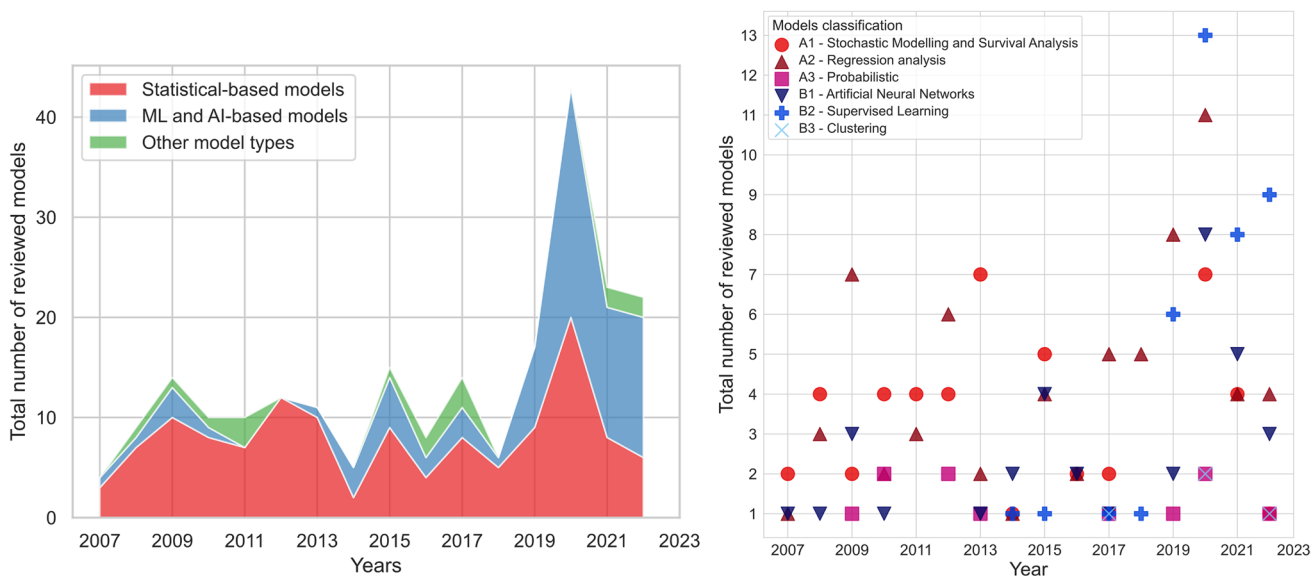**Fig. 6** Comparison of trends in DWPF model types between 2007 and 2022. The graph illustrates a substantial increase in the utilization of Machine Learning-based models in the last four years (on the left). Among these machine learning models, those employing super-

vised learning techniques have experienced a remarkable peak (on the right), surpassing models based on statistics, particularly those with regression analysis

analysis of 103 studies, the list of covariates or explanatory variables reveals the key factors influencing drinking water pipe failures, effectively addressing RQ_6. It is important to note that research in this field remains limited, primarily due to the intricacies posed by the unique local conditions of each WSDN. Consequently, there is a demand for in-depth investigations at the WSDN level to discern the impact and interrelationships among variables.

Among the noteworthy studies, Konstantinou and Stoianov (2020) stand out for their meticulous examination of explanatory variables beyond the fundamental analysis typically observed in most publications. Additionally, Robles-Velasco et al. (2020) have provided valuable insights, leading to the proposed organizational framework of variables based on different factors, as outlined in Table 7. These reference studies have significantly contributed to advancing understanding in this domain.

Kerwin et al. (2020) investigated the relationship between variables in sixteen studies, with only one study (Amaitik and Amaitik 2008) employing "time between failures" as an explanatory variable. The insights gathered from these diverse studies unequivocally demonstrate that pipes previously afflicted by failures exhibit a heightened vulnerability to subsequent failures. "Time between failures," also known as "inter-failure times", emerges as a central variable with a profound impact on the probability of failure (García-Mora et al. 2015; Le Gat 2015), attributable to its stochastic nature within the process. Thus, the inclusion of this variable becomes indispensable in forthcoming research endeavours.

The implementation of process variables varies across different studies; some opt to segregate variables based on their time dependence (Konstantinou and Stoianov 2020), while others apply all variables in one scenario and prioritize variables based on their linearity in another (Chen et al. 2019), exemplifying the importance of judiciously selecting and incorporating variables in failure modelling. It also underscores the uniqueness of failure models, necessitating bespoke approaches to suit the sensitivities of each WSDN. Such meticulous customization ensures that failure models align precisely with the complexities of the individual WSDN.

Before the advent of studies predicting the failure of drinking water pipes, the age of the pipe held paramount significance when defining renewal requirements. Counterintuitively, the wear of materials due to age does not always emerge as the most critical variable in failure processes; instead, factors such as pipe length and material composition often prove more crucial (Almheiri et al. 2020b). Several sensitivity analyses elucidate the importance of variables and their interactions within specific contexts, prompting the grouping of variables based on their relevant application environment. This categorization provides valuable insights into the criteria influencing the decision to incorporate each variable, addressing RQ_7.

Limited research has been conducted to explore the significance or impact of the identified variables on pipe failure processes, which would facilitate the correlation of these variables with risk factors associated with either inherent system functioning or external conditions. In a noteworthy study, Barton et al. (2022b) conducted a compelling analysis by estimating variable influences, confirming that the number of previous failures holds the most dominant influence, followed by pipe length and soil moisture deficit, according to the particular conditions of their study. With a different approach, Fan et al. (2021) found that time interval—related to the number of previous failures—and the ambient temperature were the most critical factors. The incidence of cold days, pipe length, and hot days as significant contributors to the pipe failure process followed this.

Prevalent explanatory variables

A fundamental objective of studies in this domain is to discern the critical explanatory variables that potentially drive variations in the frequency of failures in both the short and long terms. Furthermore, it is essential to determine how these variables can elucidate past observations and enable accurate predictions. The identified variables generally fall into three categories of paramount importance. Firstly, there are those frequently mentioned in most studies, such as the physical characteristics of the pipes, which encompass material composition, age, length, and diameter. These variables benefit from an abundance of data records provided by Water Utilities. Examining the prevalence of physical factors and their ubiquitous inclusion in nearly all studies can be readily justified by the traceability of historical and contemporary data associated with this category. Incorporating pipelines as assets within geographic information systems coupled with hydraulic models for efficient network management renders this information essential to handle for service-providing companies.

Most studies consistently highlight the substantial influence exerted by this group of variables on pipe failures (Robles-Velasco et al. 2021). The extensive impact of these variables on the modelling processes can introduce significant biases in algorithms and hinder the interpretation of results. For instance, a pertinent example is when a model indicates that a more extended pipe section correlates with a higher probability of failure (Jafari et al. 2021), linking this probability to factors such as traffic and operational stress (Mesalie et al. 2021). Numerous investigations affirm that length and the number of previous failures stand out as the most significant variables (Barton et al. 2019, 2022b; Robles-Velasco et al. 2023). It is worth contemplating the normalization of explanatory variables, including

length, during the data pre-processing stage and considering whether the normalized data adheres to a normal distribution. If not, alternative statistical distributions should be employed for standardization processes.

The number of documented previous failures serves as an operational parameter derived from the influential work of Le Gat and Eisenbeis (2000) and has demonstrated its significance in various model applications. Lastly, researchers often select additional variables tailored to the specific conditions of the studied networks. For instance, seismic activity is relevant for networks in regions susceptible to seismic events.

Physical conditions

Pipe materials constitute one of the most crucial variables influencing failures in drinking water pipes. Different materials exhibit distinct responses to changes in soil conditions and corrosion effects, necessitating the segregation of pipe material groups in failure models (Kabir et al. 2015b). As Barton et al. (2022b) stated, whether to separate or group various material types in the model inputs have been associated with differences in failure mechanisms based on material type. Nevertheless, it has been proposed that considering the influence of variables on all materials, the seasonal variation in failure rate by material, and the lower failure rate of certain materials, leading to convergence failures in the models, a more realistic approach would be to implement a global model. A global model refers to one that incorporates all materials together in the analysis. Robles-Velasco et al. (2020) indicate that a global pipeline model performs effectively, surpassing many models that segment their implementations based on material.

According to Nugroho et al. (2022) and Dawood et al. (2020b), introducing relatively new materials, such as various densities of polyvinyl chloride, and the limited availability of historical data pose challenges in the implementation of modelling compared to using data from older pipelines with different materials. These assertions stem from a survey that analysed failures in various materials for 1992 and 1993. Regardless, there may need to be more historical data for these years for plastic materials, raising questions about such claims' validity.

The mechanisms leading to pipe failures, which vary based on the material, are well-documented and can be attributed to factors such as poorly manufactured pipes, improper installation, excessive operating conditions, or third-party damage (Mohammadi and Amador Jimenez 2022; Nugroho et al. 2022). These failure mechanisms may also include susceptibility to corrosion, degradation, and structural weaknesses. Nonetheless, it is essential to acknowledge that each network's specific operational, maintenance and environmental conditions introduce variability,

preventing the generalization of observed behaviour for each material. Furthermore, the diversity in installation processes worldwide further impedes the standardization of failure causes.

An illustrative example of the variability in failure behaviour based on materials is evident when comparing results from different studies involving diverse networks and materials. For instance, Robles-Velasco et al.'s (2021) study reveals that asbestos-cement and cast-iron materials exhibit inferior performance. Conversely, Martínez García et al.'s (2021) study indicates that in comparison with asbestos-cement pipes, ductile-iron and PVC pipes exhibit higher failure rates. These findings indicate that the failure rates of specific materials can vary significantly depending on the characteristics of the case in consideration.

Several authors indicate that older pipes are expected to have significantly higher failure rates, with studies suggesting that pipe age has the most substantial influence on failure risk (Dawood et al. 2022). Regardless, in relatively young networks, age only emerges as a significant factor (Liu et al. 2022). In other cases, age is an important explanatory variable but not the most decisive in failure processes (Jafari et al. 2021; Assad and Bouferguene 2022).

According to Nugroho (2022), the precise factors influencing the relationship between pipe age and failure rate have yet to be discovered. Variations might be attributed to differences in the quality and strength of the materials used. Some studies differentiate between the age of the pipe sections and the age of the connections and fittings. This difference may be significant at an operational level since the connections may have a different age than the sections and present distinct failure patterns, such as installation errors, compared to the pipe sections, which are more influenced by other failure processes.

Only one identifies length as one of the least critical variables among the studies analysed (Almheiri et al. 2021). Most studies agree that pipe length is either the most crucial or one of the most significant variables. However, this assertion necessitates careful examination, as it is subject to essential conditioning factors. While longer pipelines are more exposed to physical risks such as loads from busy roads, varying soil and geological conditions, and an increased number of accessories due to their length, the probability of failure increases with the length, but it does not necessarily mean that length is the most decisive variable. Therefore, standardizing variables becomes a fundamental process to avoid biases in the modelling process. Nevertheless, it is essential to recognize that length remains a fundamental factor.

The relationship between pipe length and failures is significant, influenced by various physical–chemical phenomena affecting the pipe, such as the Soil Moisture Deficit (Barton et al. 2022b). These effects become more prominent

in long sections, including new installations and derivations, considering that longer pipe lengths often entail more accessories, thereby increasing the likelihood of failure (Mohammed et al. 2022). Additionally, terrain variability, factors like roads along the pipe route and pressure variations, can further influence failure probabilities over an extended pipe length.

In certain studies, researchers opted to omit pipe sections shorter than 0.5 m during data pre-processing to streamline the cleaning process of raw data (Robles-Velasco et al. 2023). Other methodologies focus on identifying short pipeline segments for specific repair and replacement, avoiding substituting simultaneously entire kilometres of pipelines (Barton et al. 2022b). However, in the dataset of some studies (Almheiri et al. 2021; Barton et al. 2022b), pipes with a minimum length of 2 m are considered, and the concept of cleaning the raw data for these shorter pipe sections is also utilized.

Environmental conditions

As previously mentioned, soil's corrosiveness is a significant environmental variable influencing water pipe integrity, particularly affecting susceptible pipe materials like steel (Kimutai et al. 2015). Soil movements and shrink–sink phenomena are considered less critical variables (Barton et al. 2022b). Regardless, these phenomena still need to be explored at the entire network level or for a more extensive pipeline dataset. Their long-term impacts under varying conditions may surpass their current understanding from laboratory-based studies. Further investigation is warranted to assess their implications in real-world scenarios comprehensively.

A notable characteristic of each WDN originates from researchers' diverse interpretations of climatic variations (Laucelli et al. 2014). A prime example is the influence of winter conditions in regions experiencing ground freezing. In such cases, it becomes essential to distinguish failure events between warm and cold seasons to avoid introducing biases (Harvey et al. 2014). Nevertheless, a comprehensive examination of the relationship between mean severity, installation, operation, maintenance conditions, and pipe failures is still needed to adequately cater to specific local requirements and address these local complexities. Further research should be directed towards addressing these aspects.

Operational and management conditions

A variable that profoundly influences pipe failure processes yet is often overlooked in most studies is the time interval between failures (Robles-Velasco et al. 2023). Whether the time between failures or since the last failure is short or long, it has various impacts. This variable is associated with discovering failure events shortly after installation or repair (Fan et al. 2021), underscoring the critical importance of Water Utilities' installation, maintenance, and pipe repair processes.

The number of previous failures also exerts a significant impact on model performance. One reason is that Water Utilities can feasibly build a failure database, providing a valuable source of case histories with sufficient traceability to rely on for data. Nevertheless, other operational factors, such as hydraulic network configuration, population density (Fan et al. 2021), pressure changes, transient phenomena, water velocity, and temperature, present complexities in the assessment and calibration phases (Robles-Velasco et al. 2020). Consequently, further research is necessary to evaluate their relationship with pipe failure processes.

A dedicated study focusing on evaluating water physico-chemical variables has revealed that residual chlorine and the number of road lanes influence the failure models' outcomes (Almheiri et al. 2021). Another significant factor is the type of soil cover surrounding the pipes, which critically impacts failure probabilities, primarily when the pipes are situated under roads with heavy traffic.

Analysis of the different model outputs

As outlined in Table 8, the outputs of the different models can be grouped according to the specific interests of each group of researchers. Nevertheless, the criteria for these groupings often need to be clarified or reported. The risk estimation approaches (Rof and CRS), widely utilized in flooding studies and climate change resilience, have yet to be utilized in this field, with only 3% of the analysed publications employing these approaches.

Conversely, the Likelihood of Failure (LoF) approach has been more widely adopted, with 35% of the studies utilizing it. As defined by Le Gat (2014), the statistical approach used for estimating failure is based on the developments proposed by Rostum (2000), which suggested the use of the non-homogeneous Poisson process (NHPP). The NHPP approach has gained acceptance within the research community due to its reliable results and solid mathematical foundation. Moreover, derivations of Rostum's processes establish a formula for calculating the probability of a model parameter given observed failure times within a specific time interval.

Failure rate (FR) is a widely used output measure, reaching 27% of use within this compendium. Statistical models have been widely used for estimating failure rates. In addition to these, as noted by Jafari et al. (2021), models such as artificial neural networks (ANNs), genetic algorithms (GAs), and fuzzy inference systems (FISs) have also

emerged as popular alternatives for modelling and predicting pipe failure rates. The successful implementation of these models in recent decades stems from their ability to capture explanatory variables' behaviour, including past failures, and the data collection improvements that enable such models' implementation.

A significant category in the model outputs reported in the literature is the number of breaks/failures (NoB), which accounts for 27% of the output variability. Use of this output had decreased over time, after reaching a peak between 2008 and 2015, when many studies employed accessible regression models to implement and interpret with commercial-available statistical software. Nonetheless, this output relies on assumptions of linear or exponential relationships between the future number of failures and some explanatory variables, such as pipe age (Karimian et al. 2021). This assumption has been challenged by several authors (Le Gat 2014), who have shown that pipe age is not a relevant predictor for some pipe materials, such as cement and plastic pipes (Robles-Velasco et al. 2020), while others have acknowledged the role of age in affecting the structural condition of the pipes (Kabir et al. 2015a), but within specific contexts and limitations.

An alternative approach to interpreting the model results involves assessing pipes estimated remaining service life from an economic standpoint. This can be achieved by comparing the equivalent annual cost of installing a new pipe with the annual cost of maintaining the existing pipe in service (Snider and McBean 2021). Additionally, deterioration curves are employed to estimate the service life, considering the concept of life cycle cost (Francisque et al. 2017). However, this output's selection criteria and analysis often need to be clarified or available in some studies (Zangenehmadar and Moselhi 2016). Previously, this output was commonly associated with a linear or exponential vulnerability increase or failure risk (Fahmy and Moselhi 2009). Nevertheless, recent advancements have shifted away from linearization by adopting alternative algorithms (Snider and McBean 2021).

Outputs, such as those based on survival probability and mean time to first/subsequent failure, are also relevant because they account for the effect of left-truncated break records (Xu and Sinha 2021). Some of these studies emphasize the number of previous failures as a key explanatory variable, showing how the selection of the modelling method influences the type of output and the explanatory variables used in the model.

Another approach to analysing the utilized outputs involves examining the relationships between the model types and their associated outcomes, as depicted in Fig. 7 for the most frequently used results. Notably, models based on supervised learning predominantly select the "Likelihood of Failure" as their output, neglecting the use of "Number of Breaks/Failures." This choice is logical, given that these algorithms' learning process involves computing probabilities for predefined categories (Jo 2021), making them ideally suited for determining likelihoods.

On the other hand, models based on regression analysis are primarily associated with the "Number of Breaks/Failures" output, as these statistical techniques aim to predict variable values based on system variables, with the number of failures being of particular interest for projection. Probabilistic models, however, do not yield a "Failure rate" output, as their focus is not on projecting future failures but on identifying failure rates concerning a risk element.
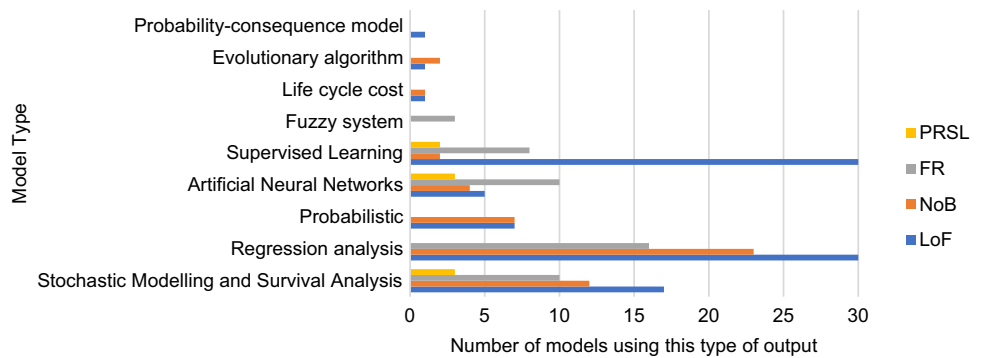
Conclusions and future research directions

This article delves into the essential field of Asset Management concerning Water Infrastructure. Managing assets is vital in addressing water infrastructure challenges, encompassing economic, health, social, and environmental aspects. Effective Infrastructure Asset Management ensures water infrastructure's long-term maintenance and adaptability, mitigating ageing effects and potential obsolescence.

Various models are utilized to predict WSDN pipe failures. These models encompass a range of approaches, including statistical models and machine learning techniques. They use historical data from Water Utilities and related explanatory variables to estimate the likelihood of future pipe failures. Some standard models used for this purpose include survival analysis models, regression models, neural networks, decision trees, and support vector machines, among others. These models play a critical role in enhancing the understanding of pipe failure dynamics and assisting Water Utilities in making informed decisions for effective asset management and infrastructure maintenance.

What criteria influence the decision to incorporate explanatory variables into the processes?

Incorporating explanatory variables in DWPF modelling processes significantly impacts these procedures. Explanatory variables provide the information that models utilize to formulate their predictions. By considering the influence of variables on predictive model accuracy and reliability, the model's robustness increases, resulting in more precise predictions. It is crucial to ensure that these data accurately represent the behaviour of the analysed system's inventories, including failure data, within a suitable timeframe that aligns with the models' algorithmic requirements and the needs defined by the service and stakeholders, subject to the existence of such data inventories.

Fig. 7 Number of selected outputs concerning the utilized model types: a graph relating the four most frequent output types

The appropriate selection of explanatory variables is essential for reducing potential interferences in understanding causal relationships and their impact on the results obtained. The use of machine learning models, which is increasingly prevalent, can complicate interpretability due to their inherent nature, making it challenging to analyse the influence of each variable on the outcome. Consequently, this hinders identifying possible improvement actions in the planning and operation/maintenance stages of Water Utilities. In complex systems like those analysed, some variables may be interrelated. For instance, some models employ the variable "time between failures" instead of "age," a technically suitable decision as it assigns more significance to data related to failure events in an unbalanced context. Furthermore, if the variables used in the models are overestimated, the quantification of uncertainty will also be affected.

Identifying the most relevant variables is critical for constructing parsimonious models and preventing overfitting, reducing computational requirements and complexities in model implementation. This consideration gains particular significance when considering the resource limitations in the context of current and future optimization efforts. The generalizability of the models for application in various scenarios and systems also relies on the choice of explanatory variables. An appropriate selection of these variables will positively influence the models' capacity to be trained and implemented on diverse datasets within the same system or for a new group with updated time ranges for the Water Utilities' inventories.

What potential future research directions exist considering DWPF modelling and explanatory variables?

Further research is required to gain a deeper understanding of the sensitivity levels of variables in the most representative models and how environmental conditions impact variable selection, enhancing our comprehension of the complex interaction between these variables. Addressing this issue necessitates executing large, controlled trials with test networks to provide more information and enhance future model accuracy.

To the best of our knowledge, no existing research has focused on exploring techniques to ensure the necessary representativeness of a selected test network. Understanding the complexity of selecting the most appropriate network for testing is crucial to improve the study of variables in pipe failure models. Assessing the impact of these variables on large distribution networks incurs high computational costs. However, by ensuring a suitable selection of a test network, these computational expenses can be significantly reduced, streamlining the optimization processes in the models.

A future approach involves identifying how the network's complexity interacts with variables and assessing whether a model suits specific network or sub-network typologies. A study's potential source of bias is the researcher's influence on the network selection, whether a section, such as a hydraulic sector, or an entire distribution network. Despite this limitation, a methodology is yet to be identified and applied to define test networks in a controlled environment to establish the proper conditions for evaluating model performance and variables.

The evidence from this study suggests that further work with a broader range of networks exhibiting more significant variability and physical location could provide deeper insights into identifying patterns or dynamics of variables and network behaviour under defined patterns of operation and maintenance. Additional research is required to examine the long-term efficacy and safety of pipe installation, operation, and maintenance procedures, as they profoundly impact the probability of failure.

An arguable weakness of the current research lies in the inability to directly apply specific models designed for one network to another without modifying the studied variables and parameters affecting each stage of the modelling process, as the methods may not be readily transferable or comparable. A standardized performance metric is crucial to interpret model results confidently. Addressing this issue requires creating and validating a methodology that allows

the evaluation of model performance, not only between different models for the same network but also between models used in different networks. This standardization requirement is vital when specifying the measurement conditions of explanatory variables, considering both their spatiotemporal variability and the range of data collection.

What limitations can be identified in the current analysed studies?

Several limitations are identified in the reviewed studies. One area for improvement is the focus on specific datasets or water supply systems, which restricts the generalizability of their findings to broader applications. Comparing results between different systems becomes challenging due to variations in data quality and time availability. Therefore, understanding how the heterogeneity of local conditions, such as material distribution, times between failures, lengths, and other essential variables, impacts the results and model fitting processes requires careful consideration.

The prevailing focus of existing studies has been on applying failure prediction models rather than conducting a comprehensive examination and understanding of the explanatory variables associated with failure processes. Most of the analysed studies do not provide a criterion to identify the most significant variables. The recent utilization of machine learning techniques has further complicated the determination of explanatory variable importance in the modelling process. Consequently, limitations in model applicability often arise due to the necessity for in-depth analysis of the variables integrated within the models.

While the current literature on DWPF primarily addresses the availability of data provided by Water Utility operators, it is crucial to emphasize the active involvement of service operators throughout the entire process, extending well beyond the mere provision of data for pre-processing purposes by researchers. The practical knowledge of operators enables modellers to comprehend the impact of variables on the models and facilitates the functional definition of hyperparameters governing the behaviour of machine learning-based models.

Ensuring adequate systems, services, and support for managing historical data, which forms the basis of the models, should be a top priority for researchers. Relying solely on blind data management at a statistical level, without considering the substantial variability inherent in local distribution networks, raises concerns about the viability of such research strategies. Expertise in modelling processes cannot replace enhancing the quality of underlying data collection processes in system administrations.

The generalizability of model outputs faces limitations that could compromise their validity. Replicating these outputs becomes challenging due to various factors, primarily when models heavily depend on Water Utility data, potentially overlooking valuable historical information. A more comprehensive approach would involve examining a large, carefully selected sample of data directly from the Water Utilities database, thereby revealing how variables interact before comparing the performance of statistical or machine-learning-based failure models.

Limited research addresses the types or groupings of characteristics significantly influencing model outcomes. Many studies solely focus on identifying the best model without thoroughly analysing local conditions and their relationship to the importance of variables or even identifying critical variables. A detailed investigation of how specific variables influence the system, or its components is necessary to determine the best model comprehensively. Therefore, exploring these variables' contextual relevance and contribution is vital to gain a deeper understanding of their impact on the system.

One aspect that has yet to be analysed in the various studies is the impact of introducing a new dataset, such as a new inventory period, into the model. Examining the results of incorporating a new dataset can help prevent overfitting, where models perform well with training data but poorly with unseen data. Additionally, bias in data collection may affect the model's ability to generalize to new scenarios. Considering these limitations can enhance the validity and applicability of future research in water pipe failure modelling and their explanatory variables.

Another crucial area for improvement is the necessity to analyse the entire universe of data a system could provide. Some studies only examine a sample that may or may not be statistically representative; however, no criteria are presented to demonstrate how this sample represents the system. By not analysing the entire system, the obtained results do not consider the system's overall impact on the model, introducing a methodological limitation. This exclusion of external behavioural factors inherent in the overall system dynamics oversimplifies the analysis, affecting the robustness of the model's results.

This study has yet to explore the time constraints of each investigation. The time horizon column of Table 7 in Appendix 1 is particularly valuable as it indicates that most of the analysed studies "predict" failures but do not specify the specific time frames of these predictions. It does not even highlight cases where the models' limitations prevent defining how many years or periods ahead the results are applicable. Making predictions for extended periods requires a sufficiently representative inventory database with a time frame like the one being predicted. Notably, in some cases, the considered base data's time for the study's execution needs to be mentioned. The age and quality of system failure inventories constrain the

development of solutions in this area. Regardless, when creating and implementing models that involve predictions, it is essential to consider the time constraints to analyse the model's implementation and the results obtained thoroughly.

What are limitations of this study?

The findings of this study offer valuable insights for researchers and water supply network planners to identify relevant variables for their pipeline failure prediction models. Yet, the scope of this document is limited to the identification of variables. It serves as a guide for the selection process, encompassing data acquisition, management, treatment, and interpretation. The unique conditions of each network and service demand a thorough examination.

The quality of the data used stands as a crucial criterion in variable selection. It has been established that the use of incomplete or low-quality data negatively impacts the predictive capabilities of the model (Fan et al. 2021). Determining data quality depends on the nature of the data and local conditions. For example, a distribution system characterized by significant pressure fluctuations, in combination with piping materials sensitive to such changes and transient events, necessitates representative temporal and spatial pressure data to incorporate pressure behaviour as an explanatory variable for failure processes.

Among the most significant limitations of this study, it was not feasible to analyse the data sample sizes, which could affect the generalizability of the conclusions to a broader population. Conducting such an analysis could suggest sample size parameters based on the type of model and variables considered. From a research perspective, we needed access to the base data of each study, preventing us from verifying the accuracy and reliability of both the base data and the results. Not knowing the characteristics of the networks being analysed further contributed to this limitation.

The examination of variable effects on the modelling processes was limited, as it relied solely on published studies, which impacted the comprehensiveness of the analysis. Additionally, not all variables could be thoroughly analysed due to space constraints, leading to selecting the most representative ones for detailed examination. The conclusions of the studies might have been influenced by external factors or unaccounted variables that were not considered in the analysis. The analysis of explanatory variables may have been more intricate, but the studies may have yet to reflect this complexity in their findings fully.

Appendix 1

See Tables 7, 8 and 9.

Factor ID descriptions are as follows:

Table 7 Explanatory variables used in modelling for predicting DWPFs

ID	Source	Model Type	Best performance model	Output Type	Time horizon (y)	P1	P2	P3	P4	P5	P6	P7	P8	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	O1	O2	O3	O4	O5	O6	O7	O8	O9	A1	A2	A3	A4
		See the references in Table 5	See Table 5	See Table 8	See Note*																																
1	Barton et al. (2022b)	M42, M43, M54	M42	LoF, CRS	–	✓	✓	✓	✓					✓		✓					✓	✓			✓									✓				
2	Nugroho et al. (2022)	M10, M15, M29, M50	M10	NoB, LCC	–		✓		✓					✓											✓												
3	Fan et al. (2021)	M17, M32, M38, M42, M45	M42	CS	–	✓	✓	✓	✓					✓		✓									✓	✓										✓	
4	Liu et al. (2022)	M17, M39	M39	LoF	–	✓	✓	✓	✓		✓				✓												✓										
5	Assad and Boufergu-ene (2022)	M17, M33, M38, M39, M42, M43, M45	M42	LoF	–	✓	✓	✓	✓				✓			✓									✓		✓										
6	Rifaai et al. (2021)	M17	M17	MtFF	–	✓	✓	✓	✓			✓				✓								✓	✓		✓							✓			
7	Almheiri et al. (2021)	M1, M2, M31, M38, M40	M40	LoF, PPA	–	✓	✓	✓	✓				✓			✓									✓		✓					✓			✓		✓
8	Robles-Velasco et al. (2021)	M32	–	FR	–	✓	✓	✓	✓		✓														✓				✓							✓	
9	Karim-ian et al. (2021)	M23	–	NoB	–	✓	✓	✓	✓					✓															✓								
10	Snider and McBean (2021)	M1, M39, M40	M40	PRSL, PRC	80	✓	✓	✓	✓			✓		✓		✓									✓	✓			✓							✓	
11	Aslani et al. (2021)	M20, M34, M39, M40, M41	M41	LoF	5	✓	✓	✓	✓							✓		✓			✓															✓	

Explanatory variables—See Table 9

Physical conditions | Environmental conditions | Operational and management conditions | Additional factors

Table 7 (continued)

ID	Source	Model Type	Best performance model	Output Type	Time horizon (y)	Explanatory variables—See Table 9																																	
						Physical conditions									Environmental conditions												Operational and management conditions									Additional factors			
		See the references in Table 5	See Table 5	See Table 8	See Note*	P1	P2	P3	P4	P5	P6	P7	P8	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	O1	O2	O3	O4	O5	O6	O7	O8	O9	A1	A2	A3	A4
12	Xu and Sinha (2021)	M1, M32	M32	MtFF	–	✓	✓	✓	✓			✓		✓																							✓		
13	Amiri-Ardakani and Najafzadeh (2021)	M20, M41, M49	M20	FR	–	✓		✓	✓				✓													✓			✓									
14	Jafari et al. (2021)	M17, M35, M37, M38, M48	M37	FR	–	✓		✓	✓				✓													✓												
15	Gouveia and Soares (2021)	M17, M28, M37, M38, M39, M42, M45	M42	CS	–	✓		✓	✓		✓															✓			✓		✓					✓		
16	Chen and Guikema (2020)	M15, M17, M18, M39, M45	M39	HBZ	5	✓		✓	✓					✓		✓				✓					✓	✓		✓							✓		
17	Weeraddana et al. (2020)	M1, M4, M15, M16, M38, M41	M1, M4, M16	SP, LoF	5	✓		✓	✓						✓	✓																		✓		✓	
18	Robles-Velasco et al. (2020)	M17, M38	M17	LoF	–	✓		✓	✓		✓														✓			✓									
19	Giraldo-González and Rodríguez (2020)	M3, M34, M38, M41, M42, M44	M42	LoF	5		✓	✓	✓		✓			✓										✓	✓	✓												

Table 7 (continued)

ID	Source	Model Type	Best performance model	Output Type	Time horizon (y)	P1	P2	P3	P4	P5	P6	P7	P8	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	O1	O2	O3	O4	O5	O6	O7	O8	O9	A1	A2	A3	A4
		See the references in Table 5	See Table 5	See Table 8	See Note*	span: Physical conditions								Environmental conditions											Operational and management conditions									Additional factors			
20	Konstantinou and Stoianov (2020)	M17, M21, M22, M28, M30, M39, M42	M42	NoB, LoF	–	✓	✓	✓	✓		✓			✓		✓									✓				✓					✓			
21	Kerwin et al. (2020)	M32	–	MtFF	15	✓	✓	✓	✓					✓	✓										✓									✓	✓		
22	Almheiri et al. (2020b)	M25, M32, M34, M43	M43	FR	40	✓	✓	✓	✓																✓												
23	Almheiri et al. (2020a)	M8	–	LoF	1											✓			✓	✓	✓				✓												
24	Elshaboury and Marzouk (2020)	M32, M34, M36, M38	M36	SCI	–	✓	✓	✓	✓	✓			✓												✓												
25	Ji et al. (2020)	M4, M10, M11	M4	LoF	–	✓	✓	✓	✓	✓		✓		✓											✓												
26	Shi et al. (2020)	M13, M32, M38, M39, M41, M45	M39	PD	–									✓		✓				✓	✓	✓												✓			
27	Gorenstein et al. (2020)	M17, M39	M39	LoF	0.5																				✓												
28	Tang et al. (2019)	M29	–	LoF	40	✓	✓	✓	✓		✓	✓		✓																				✓			
29	Sattar et al. (2019)	M33	–	NoB, FR	–	✓		✓	✓			✓													✓												
30	Chen et al. (2019)	M17, M18, M19, M39, M41	M41	LoF	–									✓	✓	✓	✓			✓	✓	✓		✓	✓										✓	✓	✓

Table 7 (continued)

ID	Source	Model Type (See the references in Table 5)	Best performance model (See Table 5)	Output Type (See Table 8)	Time horizon (y) (See Note*)	P1	P2	P3	P4	P5	P6	P7	P8	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	O1	O2	O3	O4	O5	O6	O7	O8	O9	A1	A2	A3	A4
31	Shirzad and Safari (2019)	M20, M39	M39	LoF	–		✓		✓				✓																✓								
32	Demissie et al. (2019)	M43	–	FR	20	✓	✓	✓	✓		✓	✓																									
33	Alizadeh et al. (2019)	M16, M34, M38	M16	FR	–	✓	✓	✓	✓																				✓								
34	Motiee and Ghasemnejad (2019)	M13, M14, M17	M17	NoB	–	✓	✓	✓	✓																	✓	✓		✓								
35	Kutyłowska (2019)	M38	–	FR	2	✓	✓									✓									✓												
36	Kakoudakis et al. (2018)	M23	–	LoF	1		✓	✓	✓																✓										✓		
37	Xu et al. (2018)	M14, M17	–	FR	–	✓	✓	✓	✓																✓												
38	Winkler et al. (2018)	M40	–	LoF	10	✓	✓	✓	✓																✓			✓	✓					✓	✓		
39	Chik et al. (2018)	M15	–	LoF	5												✓																				
40	Kutyłowska (2018)	M20	–	FR	1	✓	✓	✓	✓																✓												
41	Francisque et al. (2017)	M50	–	LoF, PRSL	60	✓	✓	✓	✓					✓											✓							✓					
42	Kakoudakis et al. (2017)	M23, M44	–	LoF	–		✓	✓	✓																✓												
43	Gómez-Martínez et al. (2017)	M24	–	LoF	3	✓	✓		✓				✓	✓	✓								✓		✓				✓								
44	Farmani et al. (2017)	M23, M44	–	LoF	1	✓	✓	✓	✓					✓		✓																					
45	Ward et al. (2017)	M4, M46	–	LoF	3	✓	✓	✓	✓				✓																								

Table 7 (continued)

ID	Source	Model Type (See the references in Table 5)	Best performance model (See Table 5)	Output Type (See Table 8)	Time horizon (y) (See Note*)	P1	P2	P3	P4	P5	P6	P7	P8	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	O1	O2	O3	O4	O5	O6	O7	O8	O9	A1	A2	A3	A4
46	Malinowska (2017)	M48, M49	–	RoFp	–	✓	✓						✓																								
47	Kamiński et al. (2017)	M32	–	NoB	4	✓	✓		✓					✓											✓												
48	Chik et al. (2017)	M9, M14, M29	–	LoF	1		✓																		✓				✓								
49	Sattar et al. (2016)	M49	–	MtFF, NoB	4	✓	✓	✓	✓			✓		✓											✓									✓			
50	Kabir et al. (2016)	M1, M15	–	LoF	60	✓	✓	✓	✓		✓			✓		✓																					
51	Shin et al. (2016)	M1	–	SP	–	✓	✓	✓	✓																✓												
52	Zangenehmadar and Moselhi (2016)	M32	–	PRSL	–	✓	✓	✓	✓																✓									✓			
53	Markose and Deka (2016)	M13, M34, M48	M34	NoB	1	✓	✓	✓	✓					✓											✓									✓			
54	Le Gat (2015)	M5	–	FR, SP	40	✓	✓	✓	✓		✓	✓																	✓								
55	Kimutai et al. (2015)	M1, M2, M15	M1	NoB	–	✓	✓	✓	✓	✓		✓		✓		✓				✓	✓																
56	García-Mora et al. (2015)	M17	–	LoF, RC	50	✓	✓	✓	✓						✓												✓										
57	Aydogdu and Firat (2015)	M34, M35, M36, M37	–	FR	–	✓	✓	✓	✓					✓											✓												
58	Arsénio et al. (2015)	M52	–	RPM	–	✓	✓	✓	✓					✓		✓									✓												
59	Kabir et al. (2015b)	M13, M24	–	FR	–		✓	✓	✓					✓		✓				✓	✓			✓	✓												
60	Jenkins et al. (2015)	M1, M11	–	NoB	–	✓	✓							✓											✓												

Table 7 (continued)

ID	Source	Model Type (See the references in Table 5)	Best performance model (See Table 5)	Output Type (See Table 8)	Time horizon (y) (See Note*)	P1	P2	P3	P4	P5	P6	P7	P8	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	O1	O2	O3	O4	O5	O6	O7	O8	O9	A1	A2	A3	A4
61	Kutyłowska (2015)	M32	–	FR	6	✓	✓	✓	✓		✓														✓												
62	Laucelli et al. (2014)	M23	–	LoF	2	✓		✓	✓							✓												✓									
63	Harvey et al. (2014)	M32	–	NoB	3	✓		✓	✓			✓		✓											✓										✓		
64	Shirzad et al. (2014)	M32, M38	–	FR	–			✓	✓				✓	✓											✓			✓									
65	Francis et al. (2014)	M3	–	CS	–									✓		✓	✓				✓			✓	✓												✓
66	Le Gat et al. (2013)	M4, M7, M12	–	FR, SP	15	✓		✓	✓																✓												
67	Martins et al. (2013)	M4, M6, M14	–	NoB, LoF	–	✓		✓	✓																✓												
68	Asnaashari et al. (2013)	M17, M34	–	FR	10	✓		✓	✓					✓											✓										✓		
69	Scheidegger et al. (2013)	M4, M11, M29	–	NoB, LoF	–		✓	✓																	✓												
70	Renaud et al. (2012)	M6	–	NoB	10	✓		✓	✓					✓											✓		✓										
71	Rajani et al. (2012)	M15	–	NoB	–	✓		✓	✓				✓			✓				✓	✓				✓												
72	Malm et al. (2012)	M12	–	FR	10	✓		✓	✓					✓											✓												
73	Economou et al. (2012)	M10, M11, M15	–	NoB	–		✓	✓	✓																				✓								
74	Kleiner and Rajani (2012)	M15, M17, M27, M28	–	NoB	10	✓		✓	✓			✓				✓				✓	✓				✓												
75	Singh and Adachi (2012)	M14	–	LoF	–		✓	✓																	✓												
76	Yu et al. (2012)	M17	–	FR	–			✓	✓																✓												

Table 7 (continued)

ID	Source	Model Type (See the references in Table 5)	Best performance model (See Table 5)	Output Type (See Table 8)	Time horizon (y) (See Note*)	P1	P2	P3	P4	P5	P6	P7	P8	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	O1	O2	O3	O4	O5	O6	O7	O8	O9	A1	A2	A3	A4
																									Explanatory variables—See Table 9 (Physical conditions)													(Environmental conditions)								(Operational and management conditions)					(Additional factors)			
77	Osman and Bainbridge (2011)	M1, M12, M13, M17		LoF, FR	–	✓	✓	✓	✓				✓			✓				✓					✓										✓		
78	Park et al. (2011)	M1	–	SP, NoB	–	✓					✓													✓				✓									✓
79	St. Clair and Sinha (2011)	M49	–	PI	–	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓		✓		✓		✓			✓			✓								✓		
80	Xu et al. (2011b)	M23, M51	–	NoB, LoF	2			✓	✓																✓										✓	✓	
81	Xu et al. (2011a)	M4, M51	–	NoB	–		✓	✓	✓																✓												
82	Jafar et al. (2010)	M32	–	FR	10	✓																															
83	Christodoulou et al. (2010)	M26, M27	–	PPA, LoF	35	✓	✓		✓														✓		✓	✓									✓		
84	Fares and Zayed (2010)	M47	–	RoF	–	✓			✓			✓	✓	✓								✓			✓	✓								✓			
85	Kleiner et al. (2010)	M15	–	NoB	5	✓			✓			✓								✓					✓												
86	Debón et al. (2010)	M1, M2, M17	–	HRFB	–	✓	✓	✓	✓												✓						✓										
87	Alvisi and Franchini (2010)	M1, M4	–	NoB	4	✓	✓	✓	✓	✓		✓		✓											✓												
88	Wang et al. (2009)	M13	–	NoB	–	✓	✓		✓				✓																								
89	Asnaashari et al. (2009)	M13, M14	–	NoB	–		✓	✓	✓	✓			✓												✓			✓								✓	
90	Fahmy and Moselhi (2009)	M32, M33	–	PRSL	–	✓			✓	✓		✓		✓														✓									
91	Tabesh et al. (2009)	M17, M32, M448	–	FR	1	✓	✓	✓	✓				✓															✓									

Table 7 (continued)

ID	Source	Model Type	Best performance model	Output Type	Time horizon (y)	P1	P2	P3	P4	P5	P6	P7	P8	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	O1	O2	O3	O4	O5	O6	O7	O8	O9	A1	A2	A3	A4
		See the references in Table 5	See Table 5	See Table 8	See Note*																																
92	Savic et al. (2009)	M23	–				✓		✓																✓												
93	Dridi et al. (2009)	M4	–	LoF	10	✓	✓	✓	✓																✓												
94	Economou et al. (2009)	M15, M29	–	NoB	2	✓		✓																	✓												
95	Rogers and Grigg (2009)	M4, M15	–	NoB	25	✓	✓	✓	✓					✓											✓				✓					✓	✓	✓	✓
96	Al-Barqawi and Zayed (2008)	M53	–	WMCR	5	✓	✓		✓						✓																				✓	✓	
97	Amaitik and Amaitik (2008)	M32	–	NoB	–		✓					✓		✓			✓																					
98	Park et al. (2008)	M1	–	HRFB	–	✓	✓	✓	✓					✓											✓												
99	Davis et al. (2008)	M4, M11, M12	–	PRSL	–	✓			✓	✓			✓	✓	✓		✓												✓								
100	Moglia et al. (2008)	M4, M11, M15	–	FR	–	✓	✓	✓	✓					✓													✓	✓									
101	Achim et al. (2007)	M32	–	LoF	3	✓	✓	✓																	✓										✓		
102	Davis et al. (2007)	M4, M11	–	FR	35	✓			✓	✓																												
103	Boxall et al. (2007)	M17	–	FR	–	✓	✓	✓	✓					✓											✓										✓	✓	

Table 8 Types of outputs generated in the modelling for predicting DWPFs

LoF = likelihood/probability of failure	PPA = presumed pipe age, service life length of water pipes in years	RoF = risk of pipeline failure	FR = failure rate—pipe burst/break rate	NoB = number of breaks/failures
PRSL = predicted remaining service life	PRC = pipe replacement costs	HBZ = high breakage zones	SP = survival probabilities/curves	MtFF = mean time to first/subsequent failure
RC = reliability curves	SCI = structural condition index	RPM = replacement-prioritization maps	WMCR = water pipe condition rating	HRFB = hazard rate of the first break
PI = performance index	CS = covariates sensitivity	PD = pipe deformation	CRS = Consequence risk score	LCC = Life cycle cost curves

Table 9 Classification of explanatory variables used in modelling processes for predicting DWPFs

For physical conditions	P1 Pipe material	P2 Pipe age	P3 Pipe length	P4 Pipe diameter	P5 Pipe wall thickness	P6 Pipe connections and restrains	P7 Pipe lining and coating/protection methods	P8 Installation depth
For environmental conditions	E1 Soil type and soil corrosivity factors	E2 Pipe location	E3 Climatic condition – precipitation, temperature, season	E4 Bedding	E5 Trench backfill	E6 Thawing index	E7 Freezing index	E8 Dry index—rain deficit—soil moisture	E9 Groundwater—hydrological conditions	E10 Tree roots	E11 Land use
For operational and management conditions	O1 The number of known previous failures/breaks	O2 Time interval since the last failure on the pipe	O3 Traffic conditions and loading	O4 Seismic activity	O5 Water pressure, flow, and transients	O6 Water leakage	O7 Water quality	O8 Chlorine residual	O9 Water source
For additional factors	A1 Post-failure factors	A2 Failure event type	A3 Geospatial characteristics/network type	A4 Population datasets/number of customer

*The time horizon of the models is the maximum predicted period in years for the applicability of the considered models

Declarations

Competing interests All authors declare no financial or non-financial competing interests.

References

Achim D, Ghotb F, McManus KJ (2007) Prediction of Water Pipe Asset Life Using Neural Networks. J Infrastruct Syst 13:26–30. https://doi.org/10.1061/(ASCE)1076-0342(2007)13:1(26)

AEAS (2021) Datos del sector del agua urbana 2020 - XVI Estudio Nacional de Suministro de Agua Potable y Saneamiento en España 2020. AEAS, Madrid, Spain

Ahopelto S, Vahala R (2020) Cost-benefit analysis of leakage reduction methods in water supply networks. Water 12:195. https://doi.org/10.3390/w12010195

Al-Barqawi H, Zayed T (2008) Infrastructure management: integrated AHP/ANN model to evaluate municipal water mains' performance. J Infrastruct Syst 14:305–318. https://doi.org/10.1061/(ASCE)1076-0342(2008)14:4(305)

Alegre H, Coelho ST, Covas DIC et al (2013) A utility-tailored methodology for integrated asset management of urban water infrastructure. Water Sci Technol Water Supply 13:1444–1451. https://doi.org/10.2166/ws.2013.108

Alegre H, Baptista JM, Cabrera E Jr et al (2016) Performance indicators for water supply services, 2nd edn. IWA Publishing, London

Alizadeh Z, Yazdi J, Mohammadiun S et al (2019) Evaluation of data driven models for pipe burst prediction in urban water distribution systems. Urban Water J 16:136–145. https://doi.org/10.1080/1573062X.2019.1637004

Almheiri Z, Meguid M, Zayed T (2020a) An approach to predict the failure of water mains under climatic variations. Int J Geosynth Ground Eng 6:54. https://doi.org/10.1007/s40891-020-00237-8

Almheiri Z, Meguid M, Zayed T (2020b) Intelligent approaches for predicting failure of water mains. J Pipeline Syst Eng Pract 11:04020044. https://doi.org/10.1061/(asce)ps.1949-1204.0000485

Almheiri Z, Meguid M, Zayed T (2021) Failure modeling of water distribution pipelines using meta-learning algorithms. Water Res 205:117680. https://doi.org/10.1016/j.watres.2021.117680

Alvisi S, Franchini M (2010) Comparative analysis of two probabilistic pipe breakage models applied to a real water distribution system. Civ Eng Environ Syst 27:1–22. https://doi.org/10.1080/10286600802224064

AL-Washali T, Sharma S, Lupoja R et al (2020) Assessment of water losses in distribution networks: Methods, applications, uncertainties, and implications in intermittent supply. Resour Conserv Recycl 152:104515. https://doi.org/10.1016/j.resconrec.2019.104515

Amaitik NM, Amaitik SM (2008) Development of PCCP wire breaks prediction model using artificial neural networks. Pipelines 2008. American Society of Civil Engineers, Reston, pp 1–11

Amiri-Ardakani Y, Najafzadeh M (2021) Pipe break rate assessment while considering physical and operational factors: a methodology based on global positioning system and data-driven techniques. Water Resour Manag 35:3703–3720. https://doi.org/10.1007/s11269-021-02911-6

Ananda J (2019) Determinants of real water losses in the Australian drinking water sector. Urban Water J 16:575–583. https://doi.org/10.1080/1573062X.2019.1700288

Andreou SA, Marks DH, Clark RM (1987) A new methodology for modelling break failure patterns in deteriorating water distribution systems: Theory. Adv Water Resour 10:2–10. https://doi.org/10.1016/0309-1708(87)90002-9

Aria M, Cuccurullo C (2017) bibliometrix: an R-tool for comprehensive science mapping analysis. J Informetr 11:959–975. https://doi.org/10.1016/j.joi.2017.08.007

Arsénio AM, Dheenathayalan P, Hanssen R et al (2015) Pipe failure predictions in drinking water systems using satellite observations. Struct Infrastruct Eng 11:1102–1111. https://doi.org/10.1080/15732479.2014.938660

ASCE (2022) 2021 Report Card for America's Infrastructure: Drinking Water

Aslani B, Mohebbi S, Axthelm H (2021) Predictive analytics for water main breaks using spatiotemporal data. Urban Water J 18:433–448. https://doi.org/10.1080/1573062X.2021.1893363

Asnaashari A, McBean EA, Shahrour I, Gharabaghi B (2009) Prediction of watermain failure frequencies using multiple and Poisson regression. Water Supply 9:9–19. https://doi.org/10.2166/ws.2009.020

Asnaashari A, McBean EA, Gharabaghi B, Tutt D (2013) Forecasting watermain failure using artificial neural network modelling. Can Water Resour J 38:24–33. https://doi.org/10.1080/07011784.2013.774153

Assad A, Bouferguene A (2022) Data mining algorithms for water main condition prediction—comparative analysis. J Water Resour Plan Manag 148:1–13. https://doi.org/10.1061/(asce)wr.1943-5452.0001512

Aydogdu M, Firat M (2015) Estimation of failure rate in water distribution network using fuzzy clustering and LS-SVM methods. Water Resour Manag 29:1575–1590. https://doi.org/10.1007/s11269-014-0895-5

Barton NA, Farewell TS, Hallett SH, Acland TF (2019) 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

Barton NA, Hallett SH, Jude SR (2021) The challenges of predicting pipe failures in clean water networks: a view from current practice. Water Supply. https://doi.org/10.2166/ws.2021.255

Barton NA, Hallett SH, Jude SR, Tran TH (2022a) An evolution of statistical pipe failure models for drinking water networks: a targeted review. Water Supply 22:3784–3813. https://doi.org/10.2166/ws.2022.019

Barton NA, Hallett SH, Jude SR, Tran TH (2022b) Predicting the risk of pipe failure using gradient boosted decision trees and weighted risk analysis. NPJ Clean Water 5:22. https://doi.org/10.1038/s41545-022-00165-2

Beh EHY, Zheng F, Dandy GC et al (2017) Robust optimization of water infrastructure planning under deep uncertainty using metamodels. Environ Model Softw 93:92–105. https://doi.org/10.1016/j.envsoft.2017.03.013

Belkin M, Hsu D, Ma S, Mandal S (2019) Reconciling modern machine-learning practice and the classical bias–variance tradeoff. Proc Natl Acad Sci 116:15849–15854. https://doi.org/10.1073/pnas.1903070116

Bello O, Abu-Mahfouz A, Hamam Y et al (2019) Solving management problems in water distribution networks: a survey of approaches and mathematical models. Water 11:562. https://doi.org/10.3390/w11030562

Beuken R, Eijkman J, Savic D et al (2020) Twenty years of asset management research for Dutch drinking water utilities. Water Supply 20:2941–2950. https://doi.org/10.2166/ws.2020.179

Binder H (2014) What subject matter questions motivate the use of machine learning approaches compared to statistical models for probability prediction?: Machine learning approaches, statistical models, and subject matter questions. Biom J 56:584–587. https://doi.org/10.1002/bimj.201300218

Boxall JB, O'Hagan A, Pooladsaz S et al (2007) Estimation of burst rates in water distribution mains. Proc Inst Civ Eng - Water Manag 160:73–82. https://doi.org/10.1680/wama.2007.160.2.73

Braun M (2019) Reduced order modelling and uncertainty propagation applied to water distribution networks. UNiversité de Bordeaux, Bordeaux

Carmona-Moreno C, Crestaz E, Cimmarrusti Y et al (eds) (2021) Implementing the water–energy–food–ecosystems nexus and achieving the sustainable development goals. IWA Publishing, London

Carriço N, Ferreira B, Barreira R et al (2020) Data integration for infrastructure asset management in small to medium-sized water utilities. Water Sci Technol 82:2737–2744. https://doi.org/10.2166/wst.2020.377

Chen TYJ, Guikema SD (2020) Prediction of water main failures with the spatial clustering of breaks. Reliab Eng Syst Saf 203:107108. https://doi.org/10.1016/j.ress.2020.107108

Chen TY-J, Beekman JA, David Guikema S, Shashaani S (2019) Statistical modeling in absence of system specific data: exploratory empirical analysis for prediction of water main breaks. J Infrastruct Syst 25:04019009. https://doi.org/10.1061/(asce)is.1943-555x.0000482

Chik L, Albrecht D, Kodikara J (2017) Estimation of the short-term probability of failure in water mains. J Water Resour Plan Manag 143:04016075. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000730

Chik L, Albrecht D, Kodikara J (2018) Modeling failures in water mains using the minimum monthly antecedent precipitation index. J Water Resour Plan Manag 144:06018004. https://doi.org/10.1061/(asce)wr.1943-5452.0000926

Chini CM, Stillwell AS (2018) The state of U.S. urban water: data and the energy-water nexus. Water Resour Res 54:1796–1811

Christodoulou S, Agathokleous A, Charalambous B, Adamou A (2010) Proactive risk-based integrity assessment of water distribution networks. Water Resour Manag 24:3715–3730. https://doi.org/10.1007/s11269-010-9629-5

CIRC (2019) Canada Infrastructure Report Card 2019. Can Infrustruct 1–56

Cunningham B, LaRiviere J, Wichman CJ (2021) Clustered into control: Heterogeneous causal impacts of water infrastructure failure. Econ Inq 59:1417–1439. https://doi.org/10.1111/ecin.12975

Curt C, Tourment R, Le Gat Y, Werey C (2019) Asset management of water and sewer networks, and levees: recent approaches and current considerations. In: Life-cycle analysis and assessment in civil engineering: towards an integrated vision - proceedings of the 6th international symposium on life-cycle civil engineering, IALCCE 2018, Ghent, pp 659–666

Czako Z, Sebestyen G, Hangan A (2021) AutomaticAI—a hybrid approach for automatic artificial intelligence algorithm selection and hyperparameter tuning. Expert Syst Appl 182:115225. https://doi.org/10.1016/j.eswa.2021.115225

Davis P, Burn S, Moglia M, Gould S (2007) A physical probabilistic model to predict failure rates in buried PVC pipelines. Reliab Eng Syst Saf 92:1258–1266. https://doi.org/10.1016/j.ress.2006.08.001

Davis P, Silva DD, Marlow D et al (2008) Failure prediction and optimal scheduling of replacements in asbestos cement water pipes. J Water Supply Res Technol-Aqua 57:239–252. https://doi.org/10.2166/aqua.2008.035

Dawood T, Elwakil E, Novoa HM, Delgado JFG (2020a) Artificial intelligence for the modeling of water pipes deterioration mechanisms. Autom Constr 120:103398. https://doi.org/10.1016/j.autcon.2020.103398

Dawood T, Elwakil E, Novoa HM, Gárate Delgado JF (2020b) Water pipe failure prediction and risk models: state-of-the-art review. Can J Civ Eng 47:1117–1127. https://doi.org/10.1139/cjce-2019-0481

Dawood T, Elwakil E, Mayol Novoa H, Fernando Gárate Delgado J (2022) Watermain's failure index modeling via Monte Carlo simulation and fuzzy inference system. Eng Fail Anal 134:106100. https://doi.org/10.1016/j.engfailanal.2022.106100

de Brito MM, Almoradie A, Evers M (2019) Spatially-explicit sensitivity and uncertainty analysis in a MCDA-based flood vulnerability model. Int J Geogr Inf Sci 33:1788–1806. https://doi.org/10.1080/13658816.2019.1599120

Debón A, Carrión A, Cabrera E, Solano H (2010) Comparing risk of failure models in water supply networks using ROC curves. Reliab Eng Syst Saf 95:43–48. https://doi.org/10.1016/j.ress.2009.07.004

Demissie G, Tesfamariam S, Dibike Y, Sadiq R (2019) Pipe failure prediction with consideration of climate change. In: Encyclopedia of water. Wiley, pp 1–15

Dridi L, Mailhot A, Parizeau M, Villeneuve J-P (2009) Multiobjective approach for pipe replacement based on bayesian inference of break model parameters. J Water Resour Plan Manag 135:344–354. https://doi.org/10.1061/(ASCE)0733-9496(2009)135:5(344)

Economou T, Kapelan Z, Bailey TC (2012) On the prediction of underground water pipe failures: zero inflation and pipe-specific effects. J Hydroinformatics 14:872–883. https://doi.org/10.2166/hydro.2012.144

Economou T, Kapelan Z, Bailey T (2009) A zero-inflated bayesian model for the prediction of water pipe bursts. In: Water distribution systems analysis 2008. American Society of Civil Engineers, Reston, VA, pp 1–11

Eisenbeis P (1994) Modélisation statistique de la prévision des défaillances sur les conduites d'eau potable. Université Louis Pasteur

El-Diraby TE (2021) Water Infrastructure in the 21St century: Smart and Climate-Savvy Asset Management Policies. Residential and Civil Construction Alliance of Ontario (RCCAO), Toronto

Elshaboury N, Marzouk M (2020) Comparing machine learning models for predicting water pipelines condition. In: 2020 2nd novel intelligent and leading emerging sciences conference (NILES). IEEE, pp 134–139

ERSAR (2021) Relatório Anual dos Serviços de Águas e Resíduos em Portugal (2021). Volume 1 – Caracterização do setor de águas e resíduos. Entidade Reguladora dos Serviços de Águas e Resíduos, Lisboa

Esposito C, Landrum GA, Schneider N et al (2021) GHOST: adjusting the decision threshold to handle imbalanced data in machine learning. J Chem Inf Model 61:2623–2640. https://doi.org/10.1021/acs.jcim.1c00160

Fahmy M, Moselhi O (2009) Forecasting the remaining useful life of cast iron water mains. J Perform Constr Facil 23:269–275. https://doi.org/10.1061/(ASCE)0887-3828(2009)23:4(269)

Fahrmeir L, Tutz G (2001) Multivariate statistical modelling based on generalized linear models, 2nd edn. Springer, New York

Fan J, Han F, Liu H (2014) Challenges of Big Data analysis. Natl Sci Rev 1:293–314. https://doi.org/10.1093/nsr/nwt032

Fan X, Wang X, Zhang X, Yu PE (2021) Machine learning based water pipe failure prediction: the effects of engineering, geology, climate and socio-economic factors. Reliab Eng Syst Saf. https://doi.org/10.1016/j.ress.2021.108185

Fan X, Zhang X, Yu XB (2023) Uncertainty quantification of a deep learning model for failure rate prediction of water distribution networks. Reliab Eng Syst Saf 236:109088. https://doi.org/10.1016/j.ress.2023.109088

Fares H, Zayed T (2010) Hierarchical fuzzy expert system for risk of failure of water mains. J Pipeline Syst Eng Pract 1:53–62. https://doi.org/10.1061/(asce)ps.1949-1204.0000037

Farmani R, Kakoudakis K, Behzadian K, Butler D (2017) Pipe failure prediction in water distribution systems considering static and dynamic factors. Procedia Eng 186:117–126. https://doi.org/10.1016/j.proeng.2017.03.217

Folkman S (2018) Water main break rates in the USA and Canada: a comprehensive study

FP2E/BIPE (2019) Public Water and Wastewater Services in France

Francis RA, Guikema SD, Henneman L (2014) Bayesian Belief Networks for predicting drinking water distribution system pipe breaks. Reliab Eng Syst Saf 130:1–11. https://doi.org/10.1016/j.ress.2014.04.024

Francisque A, Tesfamariam S, Kabir G et al (2017) Water mains renewal planning framework for small to medium sized water utilities: a life cycle cost analysis approach. Urban Water J 14:493–501. https://doi.org/10.1080/1573062X.2016.1223321

Gao Y (2017) Systematic review for water network failure models and cases. University of Arkansas, Fayetteville

García-Mora B, Debón A, Santamaría C, Carrión A (2015) Modelling the failure risk for water supply networks with interval-censored data. Reliab Eng Syst Saf 144:311–318. https://doi.org/10.1016/j.ress.2015.08.003

Garzón A, Kapelan Z, Langeveld J, Taormina R (2022) Machine learning-based surrogate modeling for urban water networks: review and future research directions. Water Resour Res. https://doi.org/10.1029/2021WR031808

Giraldo-González MM, Rodríguez JP (2020) Comparison of statistical and machine learning models for pipe failure modeling in water distribution networks. Water 12:1153. https://doi.org/10.3390/w12041153

Go Associados (2021) Perdas de água 2021 (SNIS 2019): Desafios para disponibilidade hídrica e avanço da eficiência so saneamiento básico. Instituto Trata Brasil, São Paulo

Gómez-Martínez P, Cubillo F, Martín-Carrasco F, Garrote L (2017) Statistical dependence of pipe breaks on explanatory variables. Water 9:158. https://doi.org/10.3390/w9030158

Gorenstein A, Kalech M, Hanusch DF, Hassid S (2020) Pipe fault prediction for water transmission mains. Water 12:2861. https://doi.org/10.3390/w12102861

Gouveia CGN, Soares AK (2021) Water connection bursting and leaks prediction using machine learning. In: World environmental and water resources congress 2021. American Society of Civil Engineers, Reston, VA, pp 1000–1013

Güngör-Demirci G, Lee J, Keck J et al (2018) Determinants of non-revenue water for a water utility in California. J Water Supply Res Technol - AQUA 67:270–278. https://doi.org/10.2166/aqua.2018.152

Gupta A, Kulat KD (2018) A selective literature review on leak management techniques for water distribution system. Water Resour Manag 32:3247–3269. https://doi.org/10.1007/s11269-018-1985-6

Harvey R, McBean EA, Gharabaghi B (2014) Predicting the timing of water main failure using artificial neural networks. J Water Resour Plan Manag 140:425–434. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000354

Herstein LM, Filion YR (2011) Closure to "Evaluating Environmental Impact in Water Distribution System Design" by L. M. Herstein, Y. R. Filion, and K. R. Hall. J Infrastruct Syst 15(17):52–53. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000045

Herstein LM, Filion YR, Hall KR (2009) Evaluating environmental impact in water distribution system design. J Infrastruct Syst 15:241–250. https://doi.org/10.1061/(ASCE)1076-0342(2009)15:3(241)

ICAMLDA (2010) Advances in machine learning and data analysis, 1st edn. Springer, Dordrecht

ISO/TC 251 (2014) ISO 55000:2014. Asset management - Overview, principles and terminology

Istituto Nazionale di Statistica (2021) Le statistiche dell'istat sull'acqua, anni 2018-2020

Jafar R, Shahrour I, Juran I (2010) Application of Artificial Neural Networks (ANN) to model the failure of urban water mains. Math Comput Model 51:1170–1180. https://doi.org/10.1016/j.mcm.2009.12.033

Jafari SM, Zahiri AR, Bozorg Hadad O, Mohammad Rezapour Tabari M (2021) 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

Jenkins L, Gokhale S, McDonald M (2015) Comparison of pipeline failure prediction models for water distribution networks with uncertain and limited data. J Pipeline Syst Eng Pract 6:04014012. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000181

Ji J, Hong Lai J, Fu G et al (2020) Probabilistic failure investigation of small diameter cast iron pipelines for water distribution. Eng Fail Anal 108:104239. https://doi.org/10.1016/j.engfailanal.2019.104239

Jo T (2021) Machine learning foundations: supervised, unsupervised, and advanced learning. Springer, Cham

Kabir G, Tesfamariam S, Francisque A, Sadiq R (2015a) Evaluating risk of water mains failure using a Bayesian belief network model. Eur J Oper Res 240:220–234. https://doi.org/10.1016/j.ejor.2014.06.033

Kabir G, Tesfamariam S, Loeppky J, Sadiq R (2015b) Integrating Bayesian linear regression with ordered weighted averaging: uncertainty analysis for predicting water main failures. ASCE-ASME J Risk Uncertain Eng Syst Part Civ Eng 1:04015007. https://doi.org/10.1061/AJRUA6.0000820

Kabir G, Tesfamariam S, Loeppky J, Sadiq R (2016) Predicting water main failures: a Bayesian model updating approach. Knowl-Based Syst 110:144–156. https://doi.org/10.1016/j.knosys.2016.07.024

Kakoudakis K, Behzadian K, Farmani R, Butler D (2017) Pipeline failure prediction in water distribution networks using evolutionary polynomial regression combined with K-means clustering. Urban Water J 14:737–742. https://doi.org/10.1080/1573062X.2016.1253755

Kakoudakis K, Farmani R, Butler D (2018) Pipeline failure prediction in water distribution networks using weather conditions as explanatory factors. J Hydroinformatics 20:1191–1200. https://doi.org/10.2166/hydro.2018.152

Kamiński K, Kamiński W, Mizerski T (2017) Application of artificial neural networks to the technical condition assessment of water supply systems. Ecol Chem Eng S 24:31–40. https://doi.org/10.1515/eces-2017-0003

Karimian F, Kaddoura K, Zayed T et al (2021) Prediction of breaks in municipal drinking water linear assets. J Pipeline Syst Eng Pract 12:04020060. https://doi.org/10.1061/(asce)ps.1949-1204.0000511

Kaur H, Pannu HS, Malhi AK (2020) A systematic review on imbalanced data challenges in machine learning: applications and solutions. ACM Comput Surv 52:1–36. https://doi.org/10.1145/3343440

Kerwin S, Adey BT (2021) Exploiting digitalisation to plan interventions on large water distribution networks. Infrastruct Asset Manag. https://doi.org/10.1680/jinam.20.00017

Kerwin S, Garcia de Soto B, Adey B et al (2020) Combining recorded failures and expert opinion in the development of ANN pipe failure prediction models. Sustain Resilient Infrastruct. https://doi.org/10.1080/23789689.2020.1787033

Kettler AJ, Goulter IC (1985) An analysis of pipe breakage in urban water distribution networks. Can J Civ Eng 12:286–293. https://doi.org/10.1139/l85-030

Kimutai E, Betrie G, Brander R et al (2015) Comparison of statistical models for predicting pipe failures: illustrative example with the city of calgary water main failure. J Pipeline Syst Eng Pract 6:04015005. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000196

Kitchenham BJ, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering, Version 2.3. Keele University, Keele

Kleiner Y, Rajani B (2012) Comparison of four models to rank failure likelihood of individual pipes. J Hydroinformatics 14:659–681. https://doi.org/10.2166/hydro.2011.029

Kleiner Y, Nafi A, Rajani B (2010) Planning renewal of water mains while considering deterioration, economies of scale and adjacent infrastructure. Water Supply 10:897–906. https://doi.org/10.2166/ws.2010.571

Konstantinou C, Stoianov I (2020) 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

Kutyłowska M (2015) Neural network approach for failure rate prediction. Eng Fail Anal 47:41–48. https://doi.org/10.1016/j.engfailanal.2014.10.007

Kutyłowska M (2018) Application of MARSplines method for failure rate prediction. Period Polytech Civ Eng 63:87–92. https://doi.org/10.3311/PPci.12559

Kutyłowska M (2019) Forecasting failure rate of water pipes. Water Supply 19:264–273. https://doi.org/10.2166/ws.2018.078

Large A (2015) State of the art review of models and software to manage drinking water pipes in the short and long term - Revue bibliographique des modèles pour gérer le patrimoine canalisations d'eau potable. Université de Bordeaux

Laucelli D, Rajani B, Kleiner Y, Giustolisi O (2014) Study on relationships between climate-related covariates and pipe bursts using evolutionary-based modelling. J Hydroinformatics 16:743–757. https://doi.org/10.2166/hydro.2013.082

Le Gat Y (2014) Extending the Yule process to model recurrent pipe failures in water supply networks. Urban Water J 11:617–630. https://doi.org/10.1080/1573062X.2013.783088

Le Gat Y (2015) Recurrent event modeling based on the yule process: application to water network asset management. Wiley, Hoboken

Le Gat Y, Eisenbeis P (2000) Using maintenance records to forecast failures in water networks. Urban Water 2:173–181. https://doi.org/10.1016/S1462-0758(00)00057-1

Le Gat Y, Kropp I, Poulton M (2013) Is the service life of water distribution pipelines linked to their failure rate? Water Supply 13:386–393. https://doi.org/10.2166/ws.2013.089

Le Gat Y, Curt C, Werey C et al (2023) Water infrastructure asset management: state of the art and emerging research themes. Struct Infrastruct Eng. https://doi.org/10.1080/15732479.2023.2222030

Lee S, Kim JH (2020) Quantitative measure of sustainability for water distribution systems: a comprehensive review. Sustain Switz 12:1–19. https://doi.org/10.3390/su122310093

Liu W, Wang B, Song Z (2022) Failure prediction of municipal water pipes using machine learning algorithms. Water Resour Manag 36:1271–1285. https://doi.org/10.1007/s11269-022-03080-w

Mailhot A, Pelletier G, Noël J-F, Villeneuve J-P (2000) Modeling the evolution of the structural state of water pipe networks with brief recorded pipe break histories: methodology and application. Water Resour Res 36:3053–3062. https://doi.org/10.1029/2000WR900185

Mailhot A, Poulin A, Villeneuve J-P (2003) Optimal replacement of water pipes: optimal replacement of water pipes. Water Resour Res. https://doi.org/10.1029/2002WR001904

Malinowska AA (2017) Fuzzy inference-based approach to the mining-induced pipeline failure estimation. Nat Hazards 85:621–636. https://doi.org/10.1007/s11069-016-2594-4

Malm A, Ljunggren O, Bergstedt O et al (2012) Replacement predictions for drinking water networks through historical data. Water Res 46:2149–2158. https://doi.org/10.1016/j.watres.2012.01.036

Markose LP, Deka PC (2016) ANN and ANFIS modeling of failure trend analysis in urban water distribution network, pp 255–264

Martínez García D, Lee J, Keck J (2021) Drinking water pipeline failure analysis based on spatiotemporal clustering and Poisson regression. J Pipeline Syst Eng Pract 12:05020006. https://doi.org/10.1061/(asce)ps.1949-1204.0000510

Martins A, Leitão JP, Amado C (2013) Comparative study of three stochastic models for prediction of pipe failures in water supply systems. J Infrastruct Syst 19:442–450. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000154

Mazumder RK, Salman AM, Li Y, Yu X (2021) Asset management decision support model for water distribution systems: impact of water pipe failure on road and water networks. J Water Resour Plan Manag 147:04021022. https://doi.org/10.1061/(asce)wr.1943-5452.0001365

Meijer D, Post J, van der Hoek JP et al (2021) Identifying critical elements in drinking water distribution networks using graph theory. Struct Infrastruct Eng 17:347–360. https://doi.org/10.1080/15732479.2020.1751664

Mesalie RA, Aklog D, Kifelew MS (2021) Failure assessment for drinking water distribution system in the case of Bahir Dar institute of technology, Ethiopia. Appl Water Sci 11:138. https://doi.org/10.1007/s13201-021-01465-7

Moglia M, Davis P, Burn S (2008) Strong exploration of a cast iron pipe failure model. Reliab Eng Syst Saf 93:885–896. https://doi.org/10.1016/j.ress.2007.03.033

Mohammadi A, Amador Jimenez L (2022) Asset management decision-making for infrastructure systems. Springer, Cham

Mohammed A, Kaddoura K, Zayed T et al (2022) Integrated reliability assessment model for drinking water networks: a case study of the city of London. Canada J Perform Constr Facil 36:04022039. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001747

Moher D, Liberati A, Tetzlaff J et al (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLOS Med 6:e1000097

Monfared Z, Molavi Nojumi M, Bayat A (2021) A review of water quality factors in water main failure prediction models. Water Pract Technol. https://doi.org/10.2166/wpt.2021.094

Morita M, Tung YK (2019) Uncertainty quantification of flood damage estimation for urban drainage risk management. Water Sci Technol 80:478–486. https://doi.org/10.2166/wst.2019.297

Mostafavi N, Gándara F, Hoque S (2018) Predicting water consumption from energy data: modeling the residential energy and water nexus in the integrated urban metabolism analysis tool (IUMAT). Energy Build 158:1683–1693. https://doi.org/10.1016/j.enbuild.2017.12.005

Motiee H, Ghasemnejad S (2019) Prediction of pipe failure rate in Tehran water distribution networks by applying regression models. Water Supply 19:695–702. https://doi.org/10.2166/ws.2018.137

Nair S, George B, Malano HM et al (2014) Water–energy–greenhouse gas nexus of urban water systems: Review of concepts, state-of-art and methods. Resour Conserv Recycl 89:1–10. https://doi.org/10.1016/j.resconrec.2014.05.007

National Audit Office (2020) Water supply and demand management. Dep Environ Food Rural Aff

Nishiyama M, Filion Y (2013) Review of statistical water main break prediction models. Can J Civ Eng 40:972–979. https://doi.org/10.1139/cjce-2012-0424

Nugroho W, Utomo C, Iriawan N (2022) A Bayesian pipe failure prediction for optimizing pipe renewal time in water distribution networks. Infrastructures 7:136. https://doi.org/10.3390/infrastructures7100136

OECD (2020) Financing water supply, sanitation and flood protection: challenges in EU member states and policy options. OECD Publishing, Paris

Office Français de la Biodiversité (2022) Observatoire des services public d'eau et d'assainissement. Panorama des services de leur performance en 2020

Okwori E, Pericault Y, Ugarelli R et al (2021) Data-driven asset management in urban water pipe networks: a proposed conceptual framework. J Hydroinformatics 23:1014–1029. https://doi.org/10.2166/hydro.2021.068

Osman H, Bainbridge K (2011) Comparison of statistical deterioration models for water distribution networks. J Perform Constr Facil 25:259–266. https://doi.org/10.1061/(asce)cf.1943-5509.0000157

Park S (2011) Estimating the timing of the economical replacement of water mains based on the predicted pipe break times using the proportional hazards models. Water Resour Manag 25:2509–2524. https://doi.org/10.1007/s11269-011-9823-0

Park S, Jun H, Agbenowosi N et al (2011) The proportional hazards modeling of water main failure data incorporating the time-dependent effects of covariates. Water Resour Manag 25:1–19. https://doi.org/10.1007/s11269-010-9684-y

Park S, Kim JW, Newland A et al (2008) Survival analysis of water distribution pipe failure data using the proportional hazards model. In: World environmental and water resources congress 2008. American Society of Civil Engineers, Reston, VA, pp 1–10

Pathirana A, Den Heijer F, Sayers PB (2021) Water infrastructure asset management is evolving. Infrastructures 6:90. https://doi.org/10.3390/infrastructures6060090

Pelletier G (2000) Impact du remplacement des conduites d'aqueduc sur le nombre annuel de bris. Universiré du Québec

Rahman S, Devera J, Reynolds J, et al (2014) Risk assessment model for pipe rehabilitation and replacement in a water distribution system. 10

Rajani B, Kleiner Y, Sink J-E (2012) Exploration of the relationship between water main breaks and temperature covariates. Urban Water J 9:67–84. https://doi.org/10.1080/1573062X.2011.630093

Renaud E, Le Gat Y, Poulton M (2012) Using a break prediction model for drinking water networks asset management: from research to practice. Water Sci Technol Water Supply 12:674–682. https://doi.org/10.2166/ws.2012.040

Rifaai TM, Abokifa AA, Sela L (2021) Integrated approach for pipe failure prediction and condition scoring in water infrastructure systems. Reliab Eng Syst Saf 220:108271. https://doi.org/10.1016/j.ress.2021.108271

Robles-Velasco A, Cortés P, Muñuzuri J, Onieva L (2020) 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.ress.2019.106754

Robles-Velasco A, Ramos-Salgado C, Muñuzuri J, Cortés P (2021) Artificial neural networks to forecast failures in water supply pipes. Sustainability 13:8226. https://doi.org/10.3390/su13158226

Robles-Velasco A, Cortés P, Muñuzuri J, De Baets B (2023) Prediction of pipe failures in water supply networks for longer time periods through multi-label classification. Expert Syst Appl 213:119050. https://doi.org/10.1016/j.eswa.2022.119050

Rogers PD, Grigg NS (2009) Failure assessment modeling to prioritize water pipe renewal: two case studies. J Infrastruct Syst 15:162–171. https://doi.org/10.1061/(ASCE)1076-0342(2009)15:3(162)

Roigé N, Pujadas P, Cardús J, Aguado A (2020) Water network renewal strategy: a case study of Aigües De Barcelona. Proc Inst Civ Eng Water Manag 173:59–68. https://doi.org/10.1680/jwama.18.00100

Rostum J (2000) Statistical modelling of pipe failures in water networks. Nor Univ Sci Technol 1–132

Rulleau B, Salles D, Gilbert D et al (2020) Crafting futures together: scenarios for water infrastructure asset management in a context of global change. Water Sci Technol Water Supply 20:3052–3067. https://doi.org/10.2166/ws.2020.185

Sattar AMA, Gharabaghi B, McBean EA (2016) Prediction of timing of watermain failure using gene expression models. Water Resour Manag 30:1635–1651. https://doi.org/10.1007/s11269-016-1241-x

Sattar AMA, Ertuğrul ÖF, Gharabaghi B et al (2019) Extreme learning machine model for water network management. Neural Comput Appl 31:157–169. https://doi.org/10.1007/s00521-017-2987-7

Savic DA, Giustolisi O, Laucelli D (2009) Asset deterioration analysis using multi-utility data and multi-objective data mining. J Hydroinformatics 11:211–224. https://doi.org/10.2166/hydro.2009.019

Scheidegger A, Scholten L, Maurer M, Reichert P (2013) Extension of pipe failure models to consider the absence of data from replaced pipes. Water Res 47:3696–3705. https://doi.org/10.1016/j.watres.2013.04.017

Scheidegger A, Leitão JP, Scholten L (2015) Statistical failure models for water distribution pipes—a review from a unified perspective. Water Res 83:237–247. https://doi.org/10.1016/j.watres.2015.06.027

Shamir U, Howard CDD (1979) An analytic approach to scheduling pipe replacement. J Am Water Works Assoc 71:248–258. https://doi.org/10.1002/j.1551-8833.1979.tb04345.x

Shi F, Peng X, Liu Z et al (2020) A data-driven approach for pipe deformation prediction based on soil properties and weather conditions. Sustain Cities Soc 55:102012. https://doi.org/10.1016/j.scs.2019.102012

Shin H, Kobayashi K, Koo J, Do M (2016) Estimating burst probability of water pipelines with a competing hazard model. J Hydroinformatics 18:126–135. https://doi.org/10.2166/hydro.2015.016

Shirzad A, Safari MJS (2019) Pipe failure rate prediction in water distribution networks using multivariate adaptive regression splines and random forest techniques. Urban Water J 16:653–661. https://doi.org/10.1080/1573062X.2020.1713384

Shirzad A, Tabesh M, Farmani R (2014) A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks. KSCE J Civ Eng 18:941–948. https://doi.org/10.1007/s12205-014-0537-8

Singh A, Adachi S (2012) Expectation analysis of the probability of failure for water supply pipes. J Pipeline Syst Eng Pract 3:36–46. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000094

Sitzenfrei R, Wang Q, Kapelan Z, Savić D (2020) Using complex network analysis for optimization of water distribution networks. Water Resour Res. https://doi.org/10.1029/2020WR027929

Snider B, McBean EA (2020) Improving urban water security through pipe-break prediction models: machine learning or survival analysis. J Environ Eng 146:04019129. https://doi.org/10.1061/(asce)ee.1943-7870.0001657

Snider B, McBean EA (2021) Combining machine learning and survival statistics to predict remaining service life of watermains. J Infrastruct Syst 27:04021019. https://doi.org/10.1061/(asce)is.1943-555x.0000629

Snyder H (2019) Literature review as a research methodology: an overview and guidelines. J Bus Res 104:333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

St. Clair AM, Sinha SK (2011) Development and the comparison of a weighted factor and fuzzy inference model for performance prediction of metallic water pipelines. In: Pipelines 2011. American Society of Civil Engineers, Reston, VA, pp 24–32

St. Clair AM, Sinha S (2012) State-of-the-technology review on water pipe condition, deterioration and failure rate prediction models! Urban Water J 9:85–112. https://doi.org/10.1080/1573062X.2011.644566

Tabesh M, Soltani J, Farmani R, Savic D (2009) Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling. J Hydroinformatics 11:1–17. https://doi.org/10.2166/hydro.2009.008

Tang K, Parsons DJ, Jude S (2019) Comparison of automatic and guided learning for Bayesian networks to analyse pipe failures in the water distribution system. Reliab Eng Syst Saf 186:24–36. https://doi.org/10.1016/j.ress.2019.02.001

Ugarelli R, Sægrov S (2022) Infrastructure asset management: historic and future perspective for tools, risk assessment, and digitalization for competence building. Water 14:1236. https://doi.org/10.3390/w14081236

Ulusoy A-J, Pecci F, Stoianov I (2021) Bi-objective design-for-control of water distribution networks with global bounds. Optim Eng. https://doi.org/10.1007/s11081-021-09598-z

Vieira J, Cabral M, Almeida N et al (2020) Novel methodology for efficiency-based long-term investment planning in water infrastructures. Struct Infrastruct Eng 16:1654–1668. https://doi.org/10.1080/15732479.2020.1722715

Wang Y, Zayed T, Moselhi O (2009) Prediction models for annual break rates of water mains. J Perform Constr Facil 23:47–54. https://doi.org/10.1061/(ASCE)0887-3828(2009)23:1(47)

Ward B, Selby A, Gee S, Savic D (2017) Deterioration modelling of small-diameter water pipes under limited data availability. Urban Water J 14:743–749. https://doi.org/10.1080/1573062X.2016.1254252

Water and Sanitation Program WB (2021) IB-NET Database - 6.1 - Non Revenue Water

Weeraddana D, Hapuarachchi H, Kumarapperuma L, et al (2020) Long-term water pipe condition assessment: a semiparametric model using gaussian process and survival analysis. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, pp 487–499

Wilson D, Filion Y, Moore I (2017) 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

Winkler D, Haltmeier M, Kleidorfer M et al (2018) 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

Xu H, Sinha SK (2021) Modeling pipe break data using survival analysis with machine learning imputation methods. J Perform Constr Facil 35:04021071. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001649

Xu Q, Chen Q, Li W (2011a) Application of genetic programming to modeling pipe failures in water distribution systems. J Hydroinformatics 13:419–428. https://doi.org/10.2166/hydro.2010.189

Xu Q, Chen Q, Li W, Ma J (2011b) Pipe break prediction based on evolutionary data-driven methods with brief recorded data. Reliab Eng Syst Saf 96:942–948. https://doi.org/10.1016/j.ress.2011.03.010

Xu Q, Qiang Z, Chen Q et al (2018) A superposed model for the pipe failure assessment of water distribution networks and uncertainty analysis: a case study. Water Resour Manag 32:1713–1723. https://doi.org/10.1007/s11269-017-1899-8

Yu PJ, Nie JX, Xu G et al (2012) Study of failure rate model for a large-scale water supply network in Southern China based on different diameters. Appl Mech Mater 260–261:1200–1205. https://doi.org/10.4028/www.scientific.net/AMM.260-261.1200

Zamenian H, Choi J, Sadeghi SA, Naderpajouh N (2017) Systematic approach for asset management of urban water pipeline infrastructure systems. Built Environ Proj Asset Manag 7:506–517. https://doi.org/10.1108/BEPAM-01-2017-0005

Zangenehmadar Z, Moselhi O (2016) Assessment of remaining useful life of pipelines using different artificial neural networks models. J Perform Constr Facil 30:04016032. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000886