

CAUSAL INFERENCE IN WATER DISTRIBUTION NETWORKS TO QUANTIFY THE EFFECTS OF NETWORK DAMAGE

School of Computer Science & Applied Mathematics
University of the Witwatersrand

Katlego L. Rammutloa
1103519

Supervised by
Mr. Rudzani Mulaudzi and,
Prof. Ritesh Ajoodha

May 17, 2025



ExplainableAI Lab
— MODELLING. DECISION MAKING. CAUSALITY —

A dissertation submitted to the Faculty of Science, University of the Witwatersrand,
Johannesburg, in partial fulfilment of the requirements for the degree of MSc Data Science.

Abstract

Water Distribution Networks (WDNs) are engineered systems of interconnected pipes, pumps, and reservoirs that deliver potable water from treatment plants to consumers. These networks are critical to public health but are highly vulnerable to structural damage (e.g., leaks, pipe corrosion), which disrupts water flow and complicates impact prediction. Current methods for assessing damage—such as hydraulic simulations and machine learning—rely on statistical correlations or optimisation, failing to model causal relationships. This limits their ability to predict cascading effects or guide repairs under uncertainty.

This study addresses these limitations by applying a causal inference framework for analysing WDNs. The framework leverages graphical causal models to represent the network's structure and quantifies the impact of damage on water flow predictions. Using Average Treatment Effect (ATE) and Mean Squared Error (MSE) metrics, we analyse how structural damage affects prediction accuracy across different network regions. The framework focuses on three critical areas: source nodes (reservoirs and entry points), mid-network nodes (junction points and main distribution pipes), and consumer nodes (end-user connection points).

Experiments on a simulated WDN reveal that damage affecting 40% or more of the network significantly compromises predictive accuracy. Mid-network and consumer nodes prove particularly vulnerable, with damage to these locations causing the greatest disruption to flow predictions. In contrast, source nodes demonstrate greater resilience due to built-in redundancies. Additionally, the study finds that treatment locations closer to outcome variables maintain predictive accuracy longer under damage conditions.

By integrating causal inference into WDN analysis, this research provides network operators with a robust methodology for evaluating damage impacts and offers actionable insights for improving network resilience. The findings contribute to both infrastructure management practices and the broader application of causal inference to complex systems analysis.

Keywords: Water distribution networks, graphical causal models, leak detection, structural damage, causal inference.

Declaration

I, Rammutloa KL (1103519), hereby declare the contents of this research proposal to be my own work. This proposal is submitted for the degree of MSc Data Science at the University of the Witwatersrand. This work has not been submitted to any other university, or for any other degree.

A handwritten signature in black ink, appearing to read 'R. Hloa.', with a stylized flourish.

May 17, 2025

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Ritesh and Rudzani, for their unwavering support and guidance throughout this journey. Their expertise, patience, and encouragement were invaluable in helping me navigate the complexities of this research.

I am also immensely grateful to the Explainable AI Lab, which provided the foundational ideas and inspiration for this study. The lab's dedication to advancing explainability and graphical models shaped my understanding and allowed me to conceptualise this research in meaningful ways.

To my parents, I extend my heartfelt thanks for instilling in me the values of discipline and focus. Their steadfast belief in my abilities has always been a source of strength and motivation.

Finally, I would like to thank my girlfriend, Zinnia, for her constant encouragement and emotional support. Her presence and understanding through every high and low of this journey have been an anchor, and I am forever grateful for her unwavering belief in me.

This research would not have been possible without the support of all these incredible individuals, to whom I owe my deepest thanks.

Contents

Preface	
Abstract	i
Declaration	ii
Acknowledgements	iii
Table of Contents	iv
List of Figures	vi
List of Tables	1
1 Introduction	1
2 Background	4
2.1 Bayesian Networks	4
2.2 Causal Models	5
2.2.1 Causal Theory	6
2.2.2 Representing Interventions	6
2.2.3 Identification	7
2.2.4 Intervention Calculus	7
3 Related Work	9
3.0.1 Conceptual Framework	9
3.0.2 Water Distribution Networks (WDNs)	10
3.0.3 Pipe Damage: A Critical Challenge in WDN Management	10
3.0.4 Current Approaches to WDN Analysis	11
3.0.5 Machine Learning and Data-Driven Approaches	12
3.0.6 Literature Gap	13
4 Methodology	15
4.1 Problem Formulation	15
4.1.1 Research Questions & Experimental Setup	16
4.2 Methodology	16
4.2.1 Causal Inference: Structural Damage	18
4.2.2 Evaluation	19
4.3 Ethical Considerations	20
4.4 Limitations	20
5 Results & Discussion	22

5.1	Research Question 1: Impact of Network Damage on Prediction Accuracy	22
5.1.1	Baseline Network Analysis	22
5.1.2	Impact of Progressive Network Damage	23
5.1.3	Error Analysis and Prediction Reliability	25
5.1.4	Implications for Network Management	26
5.2	Research Question 2: Impact of Damage Location on Prediction Accuracy	26
5.2.1	Regional Vulnerability Analysis	27
5.2.2	Error Distribution Analysis	28
5.2.3	Treatment Location Effects	29
5.3	Summary of Findings and Research Contributions	32
6	Conclusion	34
	References	40

List of Figures

1.1	Water pipeline damaged as a result of roadworks - sourced from The Goan Network	1
2.1	A BN illustrating causal dependencies between variables influencing pavement slipperiness from Pearl [2014].	5
3.1	A basic illustration of causal flow between the variables of a causal graph showing how influence flows from one variable to the next in the direction of the edges. (a) illustrates causal trail, and (b) illustrates a common effect flow.	10
4.1	Typical WDN source from Ali <i>et al.</i> [1999]	17
4.2	Simplified causal graph representing core WDN components and their relationships	18
5.1	Ground truth WDN indicating the flow of water from source nodes (reservoirs) to consumption points through junction points. The treatment variable is highlighted in red, while the outcome variable is in dark blue.	23
5.2	Average Treatment Effect (ATE) vs Percentage damage for the damaged model compared to the ground truth model with the shaded area indicates the error margin	24
5.3	Percentage change in ATE vs Percentage damage for the damaged model	25
5.4	Mean Squared Error (MSE) vs damaged percentage with the shaded area indicating the error margin.	26
5.5	A comparison of the ATE vs damage percentages across the different locations (source, junctions/mid and consumer) compared to the ground truth ATE.	27
5.6	Percentage change in ATE vs Percentage damage for the damaged model across different damage locations	28
5.7	A comparison of the MSE vs damage percentages across the different locations (source, junctions/mid and consumer).	29
5.8	A plot of the ATE measured at node C_2 against varying damage to the source nodes for a variety of treatment locations	30
5.9	A plot of the ATE measured at node C_2 against varying damage to the middle nodes for a variety of treatment locations	31
5.10	A plot of the ATE measured at node C_2 against varying damage to the consumption nodes for a variety of treatment locations	31

Chapter 1

Introduction

Water distribution networks (WDN) form the backbone of urban infrastructure, providing clean water essential for public health and economic prosperity [Phan *et al.* 2019; Li and Li 2021]. These complex interconnected systems of pipes, reservoirs, and junctions serve millions of people around the world, making their reliable operation crucial for modern society. However, WDNs are increasingly vulnerable to various forms of structural damage, ranging from gradual deterioration through corrosion and ground movement to sudden failures due to construction accidents and extreme weather events (Fig. 1.1) [Phan *et al.* 2019]. When such damage occurs, it disrupts normal flow patterns within the network, potentially compromising water delivery to consumers and complicating system maintenance.



Figure 1.1: Water pipeline damaged as a result of roadworks - sourced from [The Goan Network](#)

The analysis and prediction of damage effects in WDNs present significant challenges that current methodologies struggle to address effectively. Traditional approaches fall broadly into three categories: hydraulic simulations, network topology analysis, and machine learning techniques. Hydraulic simulations, while providing detailed flow analysis, often become computationally intractable for large-scale networks and struggle to capture the cascading effects of damage [Tabesh *et al.* 2009]. Network topology methods offer insights into structural vulnerabilities but frequently overlook the dynamic aspects of water flow [Shuang *et al.* 2019]. Machine learning methods show

promise in predicting impacts but typically require extensive training data and lack interpretability in their decision-making process [Hu *et al.* 2023]. These methods often function as “black boxes”, where the internal reasoning process remains opaque to users. For WDN operators, this lack of transparency presents significant challenges when attempting to explain and justify maintenance decisions based on model predictions.

These limitations highlight a critical gap in our ability to understand and predict how damage propagates through water distribution systems. Specifically, current methods lack the capability to model the causal relationships between network components and quantify how structural changes affect water flow predictions. This gap is particularly significant when assessing the impact of partial network damage or prioritising maintenance interventions.

Causal inference frameworks offer a promising solution to these challenges [Pearl 2009b]. By explicitly modelling the cause-and-effect relationships between network components, these frameworks capture both the structural and dynamic aspects of water distribution systems. Specifically, causal models represent the network topology through directed graphs where edges indicate direct causal influences, whilst simultaneously incorporating functional relationships that quantify how changes in one component (e.g., pressure at a junction) propagate to affect others (e.g., flow rates in connected pipes). This dual representation enables the simulation of interventions and counterfactuals—asking “what if” questions about system behaviour under various damage scenarios—that purely statistical approaches cannot address.

Our research leverages graphical causal models to represent WDNs, allowing us to quantify how damage to specific components affects water flow predictions throughout the network. This approach provides several distinct advantages over traditional methods. The explicit representation of causal relationships between network components enables more accurate modelling of damage propagation. The computational efficiency of causal inference mechanisms allows for rapid assessment of various damage scenarios. The graphical representation of network dependencies improves interpretability for network operators and maintenance teams. Furthermore, the integration of both structural and hydraulic aspects of network behaviour provides a more comprehensive understanding of system dynamics.

Our study makes several significant contributions to the field. First, we show that predictive accuracy deteriorates non-linearly with network damage, identifying critical thresholds at specific damage levels. Second, we establish that the location of damage significantly influences system resilience, with mid-network and consumer nodes showing particular vulnerability compared to source nodes. Finally, we illustrate the practicality of the causal inference framework for network operators to assess and predict the impact of structural damage on water distribution systems.

Despite these advantages, the causal inference approach faces several challenges in WDN applications. The framework relies on strong assumptions about the completeness of the causal model, which may not hold in complex real-world networks with unobserved variables. Additionally, accurate parameter estimation requires high-quality

observational data, which may be sparse or noisy in practical settings. The computational complexity of causal discovery algorithms also presents scalability challenges for very large networks, potentially requiring simplifications that compromise model fidelity.

The remainder of this study is organised as follows: Chapter 2 provides a comprehensive review of the background and Chapter 3 provides the related work, establishing the theoretical foundation for our approach. Chapter 4 describes the proposed methodology, including the framework for modelling WDNs, simulating structural damage, and applying causal inference techniques. Chapter 5 presents our experimental results and discusses their implications for network management. Finally, Chapter 6 summarises our findings, discusses their broader significance, and outlines directions for future research.

Chapter 2

Background

Building upon the problem identified in the previous chapter, this section establishes the essential background required for understanding causal models. To provide a clear progression of concepts, we first introduce Bayesian networks in Section 2.1, which form the probabilistic foundation for causal modelling. We then extend these concepts to causal models in Section 2.2, showing how they enable reasoning about interventions and effects. Finally, Section ?? examines previous approaches to analysing WDNs, highlighting the limitations that motivate our causal inference approach.

2.1 Bayesian Networks

Bayesian Networks (BNs) provide a graphical framework for representing probabilistic relationships among variables [Pearl 1988; Ajoodha 2018]. Formally, a BN consists of two components: a Directed Acyclic Graph (DAG) structure \mathcal{G} and a set of parameters θ that define conditional probability distributions. In the graph structure, nodes represent random variables and directed edges encode probabilistic dependencies between these variables.

Figure 2.1 demonstrates these concepts through a simple weather-related scenario. The graph shows how different variables influence pavement slipperiness through direct and indirect relationships. For instance, the season directly affects both rain and sprinkler activation (shown by direct edges), while its influence on pavement wetness is mediated through these intermediate variables. This absence of a direct edge between season and wetness captures our understanding that seasonal changes affect pavement wetness only through their impact on rain and sprinkler use.

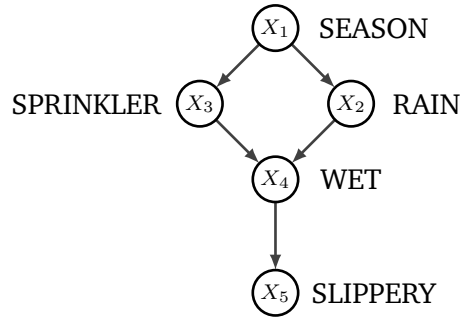


Figure 2.1: A BN illustrating causal dependencies between variables influencing pavement slipperiness from Pearl [2014].

The power of BNs lies in their ability to factorise the joint probability distribution of all variables based on local dependencies. This factorisation is expressed mathematically as:

$$P_{\mathcal{B}}(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i \mid \mathbf{Pa}_{X_i}) \quad (2.1)$$

where X_i represents the i -th variable in the network, \mathbf{Pa}_{X_i} denotes the set of parent variables that directly influence X_i , and $P(X_i \mid \mathbf{Pa}_{X_i})$ represents the conditional probability of X_i given its parents.

A fundamental concept in BNs is d-separation (directional separation), which provides a graphical criterion for identifying conditional independence relationships between variables. Consider three sets of variables X , Y , and Z in our weather example. If Z represents the wetness of the pavement, it d-separates (blocks the influence between) the season (X) and slipperiness (Y). This means that once we know the pavement's wetness, additional information about the season does not help predict slipperiness. The d-separation criterion allows us to read such independence relationships directly from the graph structure, making BNs a powerful tool for probabilistic reasoning.

2.2 Causal Models

While Bayesian networks effectively capture probabilistic relationships through the Markov condition (where each variable depends only on its parents), they do not inherently represent causation. Causation refers to the relationship between an event (the cause) and a second event (the effect), where the second event is understood as a consequence of the first. Unlike correlation, which merely indicates that two variables tend to co-occur, causation implies a directional relationship where intervention on one variable leads to changes in another Pearl [2009b]. The structure of a BN can arise from various variable orderings that satisfy the same independence patterns [Pearl 2014]. However, when we interpret a DAG as a causal model, its structure takes on additional meaning: each edge represents a direct causal relationship between variables.

Consider the weather example from Figure 2.1. As a causal model, the edge from RAIN to WET indicates that changes in rainfall directly cause changes in pavement wetness. This causal interpretation allows us to reason about interventions: what would happen to pavement wetness if we could control or manipulate rainfall? Such reasoning goes beyond the purely probabilistic relationships captured by BNs.

The transition from probabilistic to causal modelling requires a framework that can handle interventions and counterfactuals. Pearl and Verma [1995] proposed such a framework based on structural equation models (SEMs). In this framework, each variable is determined by a functional relationship with its parents:

$$X_i = f_i(\mathbf{Pa}_i, \epsilon_i) \quad (2.2)$$

where X_i is the variable of interest, f_i is a deterministic function, \mathbf{Pa}_i represents the set of parent variables that directly cause X_i , and ϵ_i is a random disturbance term representing unmodelled influences. Unlike the conditional probabilities in BNs ($P(X_i|\mathbf{Pa}_i)$), these functional relationships explicitly represent the mechanisms by which causes produce their effects.

2.2.1 Causal Theory

A complete causal theory T is formalised as a 4-tuple $T = \langle V, U, P(\mathbf{u}), \{f_i\} \rangle$, where:

- $V = \{X_1, \dots, X_n\}$ is the set of observable variables
- $U = \{U_1, \dots, U_m\}$ is the set of unobservable background variables
- $P(\mathbf{u})$ is the probability distribution over the background variables
- $\{f_i\}$ is the set of structural equations $X_i = f_i(\mathbf{Pa}_i, \mathbf{u})$

The structural equations $\{f_i\}$ together with the graph structure \mathcal{G} define how each observable variable is generated from its direct causes and background conditions [Pearl 2014]. This mechanistic representation allows us to predict not only correlations between variables but also the effects of interventions.

2.2.2 Representing Interventions

In causal models, interventions are represented by the do-operator, denoted as $do(X = x)$. This operation modifies the original causal theory T by replacing the structural equation for X with the constant x , creating a new sub-theory T_x [Pearl 2014]. For example, in our weather scenario, $do(RAIN = heavy)$ represents an intervention that sets rainfall to “heavy,” regardless of the season or other variables that normally influence rain.

The ability to represent interventions leads to three fundamental principles of causation:

1. Causation requires change: A variable X causes Y if deliberate changes in X lead to changes in Y while holding other conditions constant.
2. Causation can be potential: The causal relationship between X and Y exists even if we never observe the specific intervention.
3. Causation implies control: If X is a sufficient cause of Y , then controlling X provides a way to control Y .

The effect of an intervention is quantified by the causal effect probability $P_T(y|\hat{x})$, which represents the likelihood of observing outcome $Y = y$ when we intervene to set $X = \hat{x}$. This differs from the standard conditional probability $P(y|x)$, which represents purely observational relationships.

2.2.3 Identification

A central challenge in causal inference is determining when causal effects can be estimated from observational data alone. This is the problem of identification. Formally, a causal effect is identifiable if it can be uniquely determined from the observed probability distribution, regardless of the underlying causal mechanisms [Pearl 2009a].

More precisely, consider a class of theories $M_{\mathcal{G}}$ that share the same causal graph \mathcal{G} . A quantity Q (such as a causal effect) is identifiable in $M_{\mathcal{G}}$ if, for any two theories T_1 and T_2 in $M_{\mathcal{G}}$ that induce the same probability distribution $P_{T_1}(\mathbf{v}) = P_{T_2}(\mathbf{v})$, we have:

$$Q(T_1) = Q(T_2) \tag{2.3}$$

where \mathbf{v} represents the set of observable variables in the causal model. For causal effects specifically, we are interested in whether $P_T(y|\hat{x})$ can be computed from the observed probability distribution $P(\mathbf{v})$. This is particularly relevant when experimental manipulation is impossible or unethical, as is often the case in WDNs where we cannot deliberately damage pipes to study their effects.

2.2.4 Intervention Calculus

The ability to identify causal effects from observational data relies fundamentally on the Markov condition introduced in our discussion of Bayesian networks. In the causal context, this condition states that a variable's probability distribution is determined entirely by its direct causes (parents in the graph), with any influence from more distant ancestors flowing through these direct causes [Pearl and Verma 1995].

However, two important scenarios can complicate identification:

First, unobserved confounding occurs when an unmeasured variable influences multiple observed variables [Pearl 2014]. For example, in a WDN, an unmeasured pressure variation might affect both flow rate and pipe stress. If we observe a correlation between flow rate and pipe failures, we cannot immediately conclude causation because the unmeasured pressure could be responsible for both.

Second, selection bias arises when observations are available only for a non-random subset of the population [Pearl 2014]. In our context, this might occur if we only have data from sensors in easily accessible locations, potentially missing critical information about the network's behaviour in other areas.

To address these challenges, researchers have developed graphical rules and algebraic methods to determine when causal effects are identifiable. Do-calculus [Pearl 2014] provides three fundamental rules:

1. Rule 1 (Insertion/deletion of observations): $P(y|\hat{x}, z, w) = P(y|\hat{x}, w)$ if $(Y \perp\!\!\!\perp Z|X, W)$ in $\mathcal{G}_{\overline{X}}$
2. Rule 2 (Action/observation exchange): $P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w)$ if $(Y \perp\!\!\!\perp Z|X, W)$ in $\mathcal{G}_{\overline{XZ}}$
3. Rule 3 (Insertion/deletion of actions): $P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, w)$ if $(Y \perp\!\!\!\perp Z|X, W)$ in $\mathcal{G}_{\overline{XZ(W)}}$

Here, $\mathcal{G}_{\overline{X}}$ represents the graph obtained by removing all arrows entering X , and $(Y \perp\!\!\!\perp Z|X, W)$ denotes conditional independence of Y and Z given X and W . These rules provide a complete characterisation of identifiable relations in causal graphs.

The practical significance of these rules extends beyond their mathematical formulation. In WDNs, they help us determine when we can reliably estimate the effect of damage at one location on flow patterns elsewhere in the network, even when we cannot directly experiment with the system. This capability is crucial for developing robust prediction and maintenance strategies.

While the Markov condition provides the foundation for causal identification, the methods developed through intervention calculus allow us to handle more complex scenarios, including those with unmeasured variables and selection bias. This theoretical framework, when applied to WDNs, enables us to make reliable predictions about the effects of network damage while accounting for the practical limitations of data collection and network monitoring.

Chapter 3

Related Work

To establish the connection between causal inference and WDNs, we first present a conceptual framework that bridges these domains. We then examine current approaches to WDN analysis, including their strengths and limitations, before identifying the specific gaps our research addresses.

3.0.1 Conceptual Framework

The relationship between causal graphs and WDNs provides an intuitive framework for understanding how influence propagates through complex systems. In this context, “influence” refers to the causal impact that changes in one component have on others connected through the network structure. More formally, influence represents the quantifiable change in a downstream variable (e.g., pressure or flow rate at a consumption point) that results from an intervention on an upstream variable (e.g., valve position or pump operation). In WDNs, water flows through pipes (edges) connecting various components such as reservoirs, tanks, and junctions (nodes). Similarly, in causal graphs, influence flows through directed edges connecting variables that represent different aspects of the system [Pearl 2009b 2014].

This analogy extends beyond simple structural similarities. In a causal model, the influence of one variable on another is determined by the paths connecting them [Pearl 2009b], much like how water pressure and flow rate in one section of a WDN affect other connected sections. For instance, if we consider a variable X influencing another variable Y through intermediate variables, the causal effect follows through directed paths, similar to how water flows through a series of pipes and junctions.

The concept of d-separation provides a formal criterion for determining independence relationships between variables in a causal graph. As shown in Figure 3.1, these relationships can take two forms:

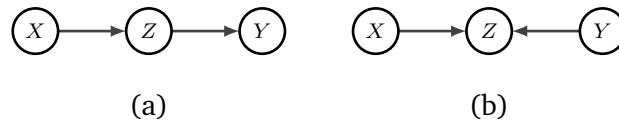


Figure 3.1: A basic illustration of causal flow between the variables of a causal graph showing how influence flows from one variable to the next in the direction of the edges. (a) illustrates causal trail, and (b) illustrates a common effect flow.

In case (a), the influence flows from X to Y through Z , representing a causal chain similar to water flowing through sequential pipes. In case (b), Z acts as a common effect of both X and Y , analogous to how multiple pipes might feed into a single junction. Understanding these patterns helps identify how interventions in one part of the network affect other components.

3.0.2 Water Distribution Networks (WDNs)

WDNs are critical infrastructure systems designed to deliver potable water from treatment plants to consumers. These complex networks comprise interconnected pipes, pumps, valves, storage tanks, and other hydraulic elements, operating under dynamic conditions to meet varying demands while maintaining adequate pressure and water quality [Mays 2000; Walski *et al.* 2003]. The complexity of WDNs arises from their extensive geographical spread, the interdependence of their components, and the need to balance multiple objectives such as water quality, system reliability, and operational efficiency [Ostfeld *et al.* 2008].

The performance of WDNs is influenced by numerous factors, including network topology, hydraulic constraints, consumer usage patterns, and environmental conditions [Ostfeld *et al.* 2008; Mays 2000]. As urban populations grow and infrastructure ages, the management and optimisation of these systems become increasingly challenging, necessitating advanced analytical approaches to ensure their reliability and resilience [Gheisi *et al.* 2016].

3.0.3 Pipe Damage: A Critical Challenge in WDN Management

Pipe damage represents one of the most significant challenges in maintaining WDN functionality. The consequences of pipe failures extend far beyond immediate operational disruptions, affecting water quality, system efficiency, and overall network resilience [Colombo and Karney 2002]. The causes of such damage are diverse, ranging from material degradation and external loads to corrosion and environmental stresses [Rajani and Kleiner 2001]. Understanding and predicting the effects of pipe damage has therefore become a central focus of research across multiple disciplines, from materials science to network analysis [Rossman 2000; Newman 2010; Shen and Osman 2021].

3.0.4 Current Approaches to WDN Analysis

Hydraulic Simulation Approaches

Hydraulic simulation has served as the traditional cornerstone of WDN analysis, providing detailed models of network behaviour under various operational scenarios. These simulations, grounded in fluid mechanics principles, solve systems of equations representing mass and energy conservation [Rossman 2000]. The evolution of simulation approaches reflects growing sophistication in handling network complexity [Kepa 2021].

Traditional demand-driven analysis (DDA) forms the foundation of hydraulic simulations, treating nodal demands as fixed while solving for pressures and flows [Todini and Pilati 1988]. However, DDA's limitations in representing low-pressure conditions during failure events led to the development of pressure-driven analysis (PDA). The PDA approach more accurately models pressure-outflow relationships, particularly crucial during pipe failures when pressure patterns deviate significantly from normal operating conditions [Giustolisi *et al.* 2008].

The application of these methods to damage analysis, as demonstrated by Tabesh *et al.* [2009], involves systematically removing pipes from network models to analyse pressure and flow distribution changes. While this approach provides detailed insights into failure consequences, its computational demands become prohibitive for large-scale networks and multiple failure scenarios. The complexity increases exponentially with network size due to the need to solve systems of non-linear equations for each potential failure configuration. The integration of Monte Carlo techniques with hydraulic simulation, pioneered by Gheisi and Naser [2014], enables more comprehensive vulnerability assessment but faces similar computational constraints for complex systems.

Recent developments in surrogate modelling and reduced-order techniques within hydraulic simulation approaches attempt to address computational efficiency while maintaining accuracy [Morosini *et al.* 2020]. However, these approaches still struggle with representing the full complexity of failure propagation mechanisms and their network-wide impacts. The primary challenges include capturing non-linear hydraulic behaviours under extreme conditions, representing transient effects during rapid changes, and accounting for the complex interactions between multiple simultaneous failures.

Graph Theory and Network Analysis

The application of graph theory to WDN analysis represents a significant advancement in understanding network vulnerability and resilience. This mathematical approach transforms physical infrastructure into abstract graphs, enabling the application of established theoretical frameworks to analyse system behaviour [Newman 2010]. The evolution of graph-theoretical approaches in WDN analysis demonstrates increasing sophistication in capturing both structural and functional aspects of network performance.

Yazdani and Jeffrey [2011] established the foundation for complex network analysis in WDNs by introducing quantitative metrics for system resilience assessment. Their

methodology employs network science concepts such as link density, clustering coefficients, and centrality measures to evaluate failure resistance characteristics. This initial framework, while innovative in its topological approach, revealed limitations in capturing the hydraulic dynamics that govern actual network behaviour. Specifically, pure graph-theoretical metrics failed to account for critical hydraulic parameters such as pressure gradients, flow directionality, and energy conservation principles that determine actual water movement through the network.

Addressing these limitations, [Di Nardo et al. \[2018\]](#) developed an integrated framework combining topological metrics with hydraulic performance indicators. Their hybrid methodology synthesizes edge connectivity analysis with flow distribution patterns, enabling more comprehensive evaluation of network robustness. The integration of algebraic connectivity measures with hydraulic parameters provides deeper insights into system behaviour under stress conditions, though challenges remain in representing temporal dynamics and non-linear responses to damage. These challenges include accurately modelling time-dependent phenomena such as pressure transients, water hammer effects, and demand pattern variations throughout daily or seasonal cycles.

Further refinements in graph-theoretical approaches emerged through spectral analysis techniques. [Farahmandfar et al. \[2020\]](#) demonstrated how eigenvalue analysis of network adjacency matrices can reveal structural vulnerabilities not apparent through traditional metrics. Their work establishes correlations between spectral properties and system resilience, though the interpretation of these mathematical relationships in terms of physical network behaviour presents ongoing challenges.

3.0.5 Machine Learning and Data-Driven Approaches

The emergence of advanced computational capabilities and big data analytics has catalysed the adoption of machine learning techniques in WDN analysis. Examples include the application of convolutional neural networks for spatial pattern recognition in pressure data, recurrent neural networks for time-series forecasting of demand fluctuations, and reinforcement learning algorithms for optimising pump schedules and valve operations [[Truong et al. 2024](#); [Zanfei et al. 2022](#); [Ma et al. 2024](#)]. These machine learning (deep learning) techniques offer new perspectives on pattern recognition and predictive modelling, complementing traditional analytical approaches [[Goodfellow et al. 2016](#)]. However, their effectiveness depends critically on data quality and quantity, presenting unique challenges in implementation.

Early applications of deep learning in WDNs focused on anomaly detection and system monitoring. [Mounce et al. \[2011\]](#) implemented artificial neural networks for real-time burst detection using pressure and flow measurements. Their approach demonstrated success in rapid event identification but revealed limitations in explaining the underlying mechanisms of failure propagation. The black-box nature of these models complicates their integration with domain expertise and physical constraints.

Significant advancement came through [Shen and Osman \[2021\]](#)'s integration of deep learning with graph neural networks. Their hybrid architecture captures both spatial relationships and topological features, enabling more accurate prediction of hydraulic

impacts following pipe failures. The model's ability to learn complex network dynamics represents a substantial improvement over traditional statistical approaches. However, the extensive data requirements and computational intensity pose challenges for practical implementation in large-scale systems.

Recent developments explore reinforcement learning for operational optimisation and transfer learning for model generalisation [Sanz *et al.* 2016; Nault and Karney 2020]. These approaches show promise in adapting to varying network configurations but face challenges in maintaining prediction reliability across different operational conditions. The integration of physical constraints and domain knowledge remains a significant challenge in ensuring model validity and interpretability.

3.0.6 Literature Gap

The synthesis of current approaches reveals several fundamental limitations in analysing pipe damage effects in WDNs. These limitations span methodological, computational, and practical dimensions, highlighting the need for more comprehensive analytical frameworks.

The integration of hydraulic behaviour with network topology remains inadequate in current methods. While hydraulic simulations provide detailed flow analysis and graph theory offers structural insights, few approaches successfully combine these perspectives. This disconnection limits our understanding of how structural damage translates into functional impacts across the network [Shuang *et al.* 2019].

Current methodologies also lack explicit representation of causal mechanisms in damage propagation. The absence of formal causal modelling frameworks impedes our ability to predict how local failures cascade through the system and affect global network performance [Zeng *et al.* 2017]. This limitation becomes particularly acute when analysing complex failure scenarios involving multiple interacting components.

Computational efficiency presents another significant challenge, particularly for large-scale networks. Detailed hydraulic simulations, while accurate, often become computationally intractable for real-time analysis and system-wide optimisation [Berardi *et al.* 2008]. This computational burden limits the practical application of these methods in operational contexts where rapid assessment and response are crucial.

The quantification of uncertainty in damage scenarios and their effects remains inadequate in existing approaches. The inherent variability in network behaviour and the complexity of failure mechanisms require sophisticated uncertainty-handling methods. Current approaches often produce deterministic predictions that fail to capture the full range of possible outcomes, potentially leading to suboptimal decision-making [Kabir *et al.* 2015].

Scalability and interpretability limitations affect many current methods, particularly those based on machine learning approaches. As networks grow in size and complexity, maintaining model accuracy while ensuring results remain interpretable becomes increasingly challenging [Hu *et al.* 2023]. This tension between model sophistication and practical utility represents a fundamental challenge in WDN analysis.

Graphical causal models offer a promising approach to address these limitations. By providing a framework that can integrate hydraulic principles with network topology while explicitly modelling causal relationships, these models have the potential to revolutionise our understanding of pipe damage effects in WDNs [Pearl 2009c].

The strengths of graphical causal models in this context include:

1. Explicit representation of causal relationships, allowing for more intuitive and interpretable models of damage propagation.
2. Improved computational efficiency by focusing on key causal pathways rather than exhaustive simulations [Spirtes *et al.* 2000].
3. Enhanced ability to quantify uncertainties through probabilistic reasoning and sensitivity analysis [Koller and Friedman 2009].
4. Facilitation of knowledge integration, allowing for the incorporation of domain expertise and prior knowledge into the modelling process [Fenton and Neil 2012].
5. Potential for improved generalisability across different network configurations and damage scenarios.

By building on the strengths of previous approaches while addressing their limitations, research into graphical causal models promises to advance our understanding of pipe damage effects in WDNs. This approach has the potential to provide new insights into the propagation of damage effects, support more effective decision-making in network management, and ultimately contribute to the development of more resilient water distribution systems.

While significant progress has been made in analysing the effects of pipe damage in WDNs, there remains ample opportunity for advancement. The development and application of graphical causal models in this domain hold promise for addressing current limitations and pushing the boundaries of our understanding of these critical infrastructure systems.

Chapter 4

Methodology

This chapter outlines the methodological approach for analysing the effects of structural damage on WDNs. Our study aims to apply a causal inference framework for predicting and understanding how damage to specific components of a WDN impacts water levels throughout the system. We begin by formulating the problem mathematically, representing the WDN as a graph structure. This formulation allows us to model the complex interactions between various network components such as reservoirs, tanks, and junctions. We then present our research questions, each accompanied by a detailed experimental setup designed to investigate the impact of network damage on water level predictions. Our methodology employs causal inference techniques to estimate the effects of interventions on the network, providing insights into the system's behaviour under various damage scenarios. Through this approach, we seek to contribute to the development of more resilient and efficient water distribution systems.

4.1 Problem Formulation

Consider a WDN consisting of reservoirs, tanks, junction points (which can be inflow or consumption points), and connecting pipes. This network can be represented as a directed graph $G = (V, E)$, where V is the set of vertices representing reservoirs, tanks, and junction points, and E is the set of edges representing pipes. Each vertex $v \in V$ has an associated water level $L(v)$.

The overarching problem that this research aims to address is: How can we model and study the effects of pipe damage on the water levels at downstream points in the network?

Mathematically, we can formulate this problem as follows: Given a graph $G = (V, E)$ representing a WDN, the water levels $L(v)$ for all $v \in V$, and a set of damaged pipes $D \subseteq E$, how can we find a function f that predicts the water level at any vertex v , given the network structure G and the set of damaged pipes D :

$$L(v) = f(G, \{L(v)|v \in V\}, D).$$

This formulation allows us to model the steady-state flow dynamics in the network while incorporating the effects of network damage on water flow. By solving this problem, we aim to understand how structural damage to the network affects water distribution and develop methods to predict and mitigate the impact of such damage.

4.1.1 Research Questions & Experimental Setup

Based on our problem formulation, we addressed two primary research questions:

Research Question 1: How does network damage, simulated by removing pipes in the WDN, impact the accuracy of predicting water levels at downstream points?

To address this question, we developed a ground truth (GT) model of a WDN using a causal graph G that models the causal relationships between its variables (components) X_1, \dots, X_n , where each variable X_j is a function f_j of its parent variables PA_j and some noise N_j . We then generated a damaged model G' by randomly removing edges from the GT model. The average treatment effect (ATE) of changing the water level at a particular junction on the water level at a consumption point downstream is measured for both G and G' . We compared the estimated ATE from the GT and the damaged models using Mean Squared Error (MSE).

The evaluation involved quantifying the difference in ATE estimates between the GT and damaged models and analysing how the extent of damage (percentage of edges removed) correlates with prediction accuracy.

Research Question 2: Does the location of the network damage (e.g., removing edges near the source vs. near consumers) influence the severity of the impact on water level prediction accuracy?

To investigate this question, we used the same GT model as in RQ1 and created multiple damaged models G' by removing edges at different locations: near the source (reservoirs), mid-network (junctions), and near consumers. For each damaged model, we estimated the ATE as in RQ1 and compared the ATEs and prediction errors across different damage locations.

The evaluation involved analysing patterns in prediction errors (MSE) across different damage locations and determining if certain areas of the network are more critical for maintaining accurate predictions.

Through these experiments, we have gained insights into the impact of network damage on water level predictions and identified critical areas within the network that have the most significant influence on prediction accuracy. This knowledge contributes to developing more robust and resilient water distribution systems and improving damage mitigation strategies.

4.2 Methodology

To perform causal inference, we need a model that captures the underlying physical processes governing water flow within the network. The causal relationships of the

nodes are governed by the graph structure that models our WDN. Figure 4.1 illustrates a graph of a typical WDN [Ali et al. 1999]. In the diagram, the flow of water between nodes is represented by edges, signifying pipes. These nodes themselves can be junctions, tanks, or reservoirs. Junctions act as connectors for two pipe segments and can also serve as points of water consumption (think residential or industrial customers) or external water sources (like a water treatment plant). Tanks and reservoirs, depicted as external nodes in the diagram, function as dedicated storage areas for water.

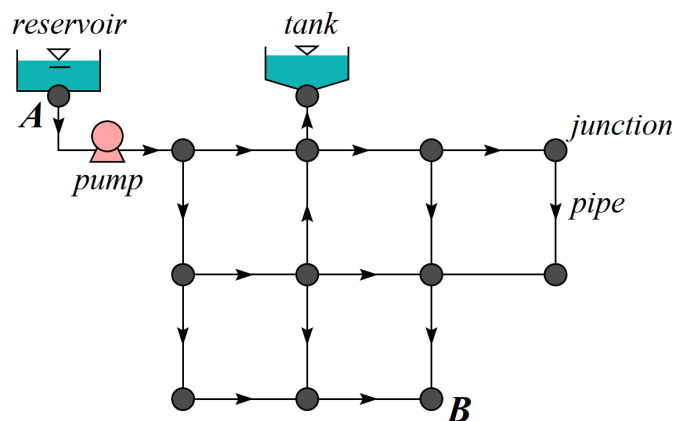


Figure 4.1: Typical WDN source from Ali et al. [1999]

Figure 4.1 illustrates a complex real-world WDN with multiple interconnected components. For our analysis, we distill this complex network into a simplified yet representative causal graph. This simplified model, shown in Figure 4.2, captures the essential causal relationships and flow dynamics present in typical WDNs while remaining tractable for detailed analysis.

In developing our methodology, we opted for a simplified WDN model rather than using pre-existing simulations or datasets for several key reasons. First, existing datasets often lack the controlled variability necessary to isolate and analyse specific causal relationships under damage conditions. By constructing our own model, we gain precise control over the causal structure, enabling systematic investigation of damage effects without confounding variables present in real-world data. Second, though simplified, our model incorporates the fundamental components and relationships that define valid WDNs according to established literature [Walski et al. 2003; Rossman 2000]. Specifically, our model includes:

1. Source nodes (reservoirs, R) that serve as water entry points with stable pressure characteristics
2. Storage components (T) that buffer supply and demand variations
3. Junction nodes (J_1, J_2) that distribute flow through the network
4. Consumption points (C_1, C_2) representing end-user demands
5. Directed connections representing pipes with associated flow capacities

This structure satisfies the minimum requirements for hydraulic network modelling whilst remaining computationally tractable for causal analysis. The simplification represents a methodological trade-off between model complexity and analytical clarity, focusing on establishing fundamental causal relationships that can later be extended to more complex configurations. Future work will validate these findings against larger, more detailed network simulations.

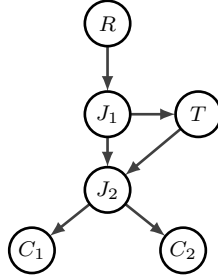


Figure 4.2: Simplified causal graph representing core WDN components and their relationships

The water level at each node is influenced by its parent nodes and associated noise factors. We express these relationships through a system of deterministic functions, employing linear functions with additive noise. Each equation represents a specific causal mechanism in the network:

$$\begin{aligned}
 R &= N_R \\
 J_1 &= \alpha_1 R + N_{J_1} \\
 T &= \alpha_2 J_1 + N_T \\
 J_2 &= \alpha_3 J_1 + \alpha_4 T + N_{J_2} \\
 C_1 &= \alpha_5 J_2 + N_{C_1} \\
 C_2 &= \alpha_6 J_2 + N_{C_2}
 \end{aligned} \tag{4.1}$$

In these equations, R represents the water level at the reservoir node, which is initialised by noise N_R . The junction nodes J_1 and J_2 represent intermediate water levels, while T denotes the water level at the tank node. The consumption points C_1 and C_2 represent the final water levels at end-user locations. The terms N_R , N_{J_1} , N_T , N_{J_2} , N_{C_1} , and N_{C_2} are independent random noise variables that represent unobserved disturbances in the system. The constants α_i (where $i \in \{1, 2, 3, 4, 5, 6\}$) represent the proportion of water flow influence from one node to another, determined by physical parameters such as pipe diameter and flow resistance.

4.2.1 Causal Inference: Structural Damage

Our approach to estimating causal effects in the network employs three fundamental identification techniques. The first technique, backdoor adjustment, controls for confounding variables that might affect both the treatment and outcome variables. In our

WDM context, this involves adjusting for upstream nodes that influence both the intervention point and the measurement point. The second technique utilises instrumental variables, which affect the treatment but not the outcome directly, helping isolate the causal effect of interest. In our network, certain upstream components can serve as instrumental variables when analysing downstream effects. The third technique, front-door adjustment, identifies causal effects by measuring the mechanism through which the treatment affects the outcome, particularly useful when analysing flow paths in the network.

We implement these techniques using *DoWhy*, a Python package specifically designed for causal inference. Causal inference refers to the process of determining the independent, actual effect of a particular phenomenon that is a component of a larger system. In our context, it involves estimating how interventions on one part of the network (such as simulated damage) affect outcomes at other locations, whilst controlling for confounding factors that might influence both the intervention and outcome variables. The package enables us to estimate the Average Treatment Effect (ATE) of interventions on our network:

$$ATE = E[C_1|do(J_1 = j')] - E[C_1|do(J_1 = j)] \quad (4.2)$$

where j' and j represent different levels of water flow from junction J_1 . Since each node represents water flow, we can express an intervention as a flow adjustment factor $\beta \in \mathbb{R}$:

$$j' = \beta j \quad (4.3)$$

4.2.2 Evaluation

Now that we can simulate network damage and estimate causal effects under damage conditions, let's explore how *DoWhy* can help us answer our specific research questions:

Research Question 1

- We first estimated the true causal effects (e.g., ATE of changing the water level at a junction on the water level at the consumer) using the ground-truth causal model.
- For each damage scenario (removing a specific edge), we defined the intervention and updated the causal model in *DoWhy*. Subsequently, we re-estimated the causal effects of interest under the damaged network conditions.
- The difference between the true effects (ground-truth model) and the estimated effects (damaged model) was quantified using metrics like MSE. This provides a measure of how network damage impacts the accuracy of causal inference.

Research Question 2

- We repeated the process outlined in the first research question for multiple damage scenarios, strategically removing edges at different locations in the network (i.e., near reservoirs, near consumption points, etc.).
- By comparing the MSE values across different damage locations, we can identify any patterns or trends. This helps us understand if the location of the damage influences the severity of the impact on causal inference accuracy.

MSE is an appropriate performance measure for this research due to several factors. MSE is particularly sensitive to large errors, which is crucial in WDN analysis where significant deviations in pressure or flow predictions could indicate critical failure points. It provides a single, easily interpretable metric that captures both the bias and variance of predictions, allowing for comprehensive comparison between different models or approaches.

By leveraging DoWhy’s functionalities for interventions, causal effect estimation, and various identification strategies, we can comprehensively analyse the impact of network damage on causal inference in WDNs. This research can provide valuable insights for network operators to improve their ability to diagnose and address issues even when the network is compromised.

4.3 Ethical Considerations

Synthetic data, by definition, does not involve any real-world personal information or identifiable details. Therefore, ethical concerns surrounding privacy and confidentiality are not applicable.

4.4 Limitations

The proposed methodology faces several significant limitations. First, our causal inference techniques rely on strong assumptions about the nature of relationships within the network. These assumptions include the constancy of causal relationships over time, the absence of hidden variables influencing both intervention and outcome, and complete observability of relevant network states. While these assumptions enable tractable analysis, they may not fully reflect real-world conditions.

Second, the accuracy of our findings depends heavily on the quality of the ground truth model and the fidelity of simulated data to real-world conditions. The physical parameters used in the model must accurately reflect actual network characteristics, which can be challenging to achieve in practice. Furthermore, our current approach models only complete blockages rather than partial damage, limiting its applicability to real-world scenarios where damage often occurs on a continuous spectrum.

To address these limitations, we propose several directions for future research. The integration of real-world sensor data would significantly improve model accuracy and validation capabilities. The development of more nuanced damage simulation capabilities, including partial blockages, would enhance the model’s practical utility. Additionally,

extending the framework to handle time-varying causal relationships and incorporating uncertainty quantification in causal effect estimates would improve its robustness. Finally, enhancing the model to consider multiple simultaneous failure modes would better reflect real-world network vulnerability scenarios.

Chapter 5

Results & Discussion

This chapter presents a comprehensive analysis of our experimental findings addressing two fundamental research questions in WDNs resilience: (1) How does network damage affect the accuracy of predicting water levels at downstream points? and (2) Does the location of network damage influence the severity of the impact on water level prediction accuracy?

Our investigation employs multiple analytical approaches, including Average Treatment Effect (ATE) analysis, Mean Squared Error (MSE) measurements, and location-specific impact assessments. The results are presented systematically, with each section building upon previous findings to provide a coherent understanding of network vulnerability and prediction reliability. Supporting visualisations are integrated throughout the discussion to illustrate key patterns and relationships discovered in the analysis.

5.1 Research Question 1: Impact of Network Damage on Prediction Accuracy

This section examines how structural damage affects the network's ability to predict water levels accurately. We present a progression of analyses that reveal critical thresholds in network damage and their implications for prediction reliability. The findings demonstrate non-linear relationships between damage extent and prediction accuracy, with several key transition points that have significant implications for network maintenance strategies.

5.1.1 Baseline Network Analysis

To establish our analytical foundation, we first constructed a ground truth model of the WDN, as shown in Figure 5.1. The selection of treatment and outcome variables was guided by network topology and practical considerations in water management. Specifically, we selected junction J_1 as our treatment variable due to its strategic position connecting multiple flow paths and its capacity to influence downstream water levels. This

junction represents a critical control point where operators could potentially intervene to regulate water flow, making it an ideal candidate for analysing causal effects.

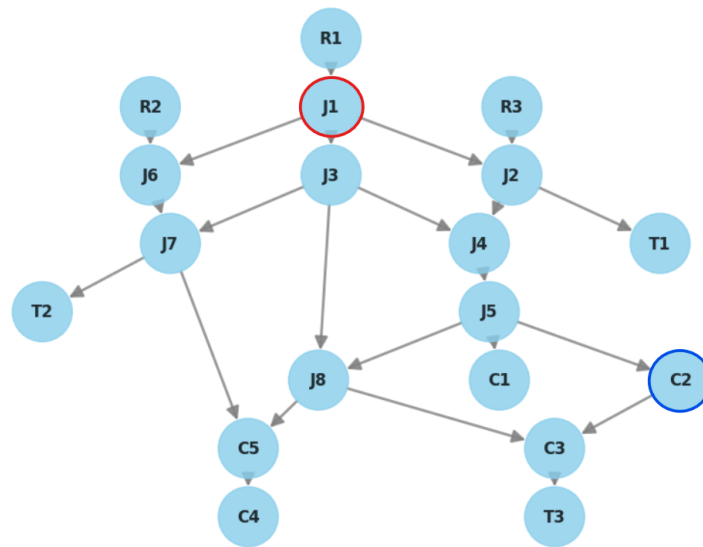


Figure 5.1: Ground truth WDN indicating the flow of water from source nodes (reservoirs) to consumption points through junction points. The treatment variable is highlighted in red, while the outcome variable is in dark blue.

The network structure reflects typical WDN architectures, with water flowing from source nodes (R1, R2 and R3) through intermediary junctions to end-user consumption points. By comparing the causal effects measured in this ground truth model against those in damaged scenarios, we can quantify how structural compromises affect our ability to predict and control water levels throughout the system.

5.1.2 Impact of Progressive Network Damage

Our analysis of the Average Treatment Effect (ATE) reveals how network damage progressively degrades our ability to predict the causal impact of interventions at our selected treatment point. Figure 5.2 illustrates this relationship by comparing the ATE values from damaged network configurations against the ground truth model's baseline ATE of 0.128 (shown by the horizontal reference line).

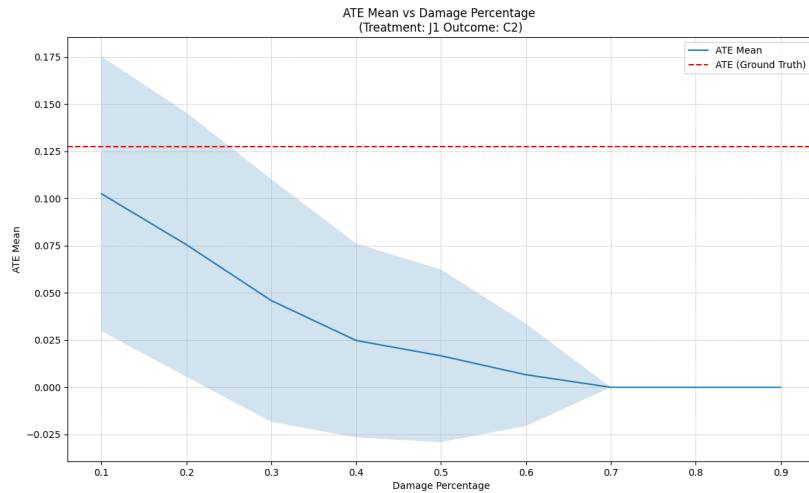


Figure 5.2: Average Treatment Effect (ATE) vs Percentage damage for the damaged model compared to the ground truth model with the shaded area indicates the error margin

The relationship between ATE and network damage reveals several critical thresholds. The network exhibits particularly high sensitivity in the early damage range, where structural changes between 10% and 30% damage trigger substantial prediction deterioration. Within this range, each 10% increase in damage results in approximately a 25% decrease in ATE, with values declining from 0.1 to 0.05. The overall non-linear relationship demonstrates that even minor structural compromises can significantly impair the network’s predictive capabilities. As damage extends beyond 70%, ATE values fall to zero, marking a practical threshold where prediction becomes unreliable.

To quantify the rate of performance degradation, Figure 5.3 examines the percentage change in ATE across damage levels.

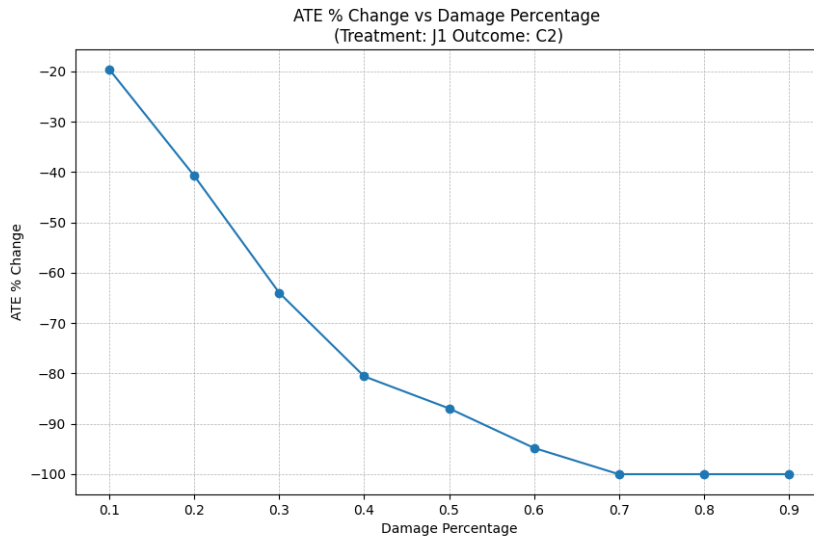


Figure 5.3: Percentage change in ATE vs Percentage damage for the damaged model

The analysis reveals an acceleration in prediction deterioration between 10% and 40% damage, with the steepest decline occurring near the 25% damage mark. This pattern identifies a critical vulnerability window where network performance rapidly degrades, suggesting the importance of maintaining network integrity below the 40% damage threshold.

5.1.3 Error Analysis and Prediction Reliability

To understand how prediction errors evolve with increasing damage, Figure 5.4 presents the Mean Squared Error (MSE) analysis.

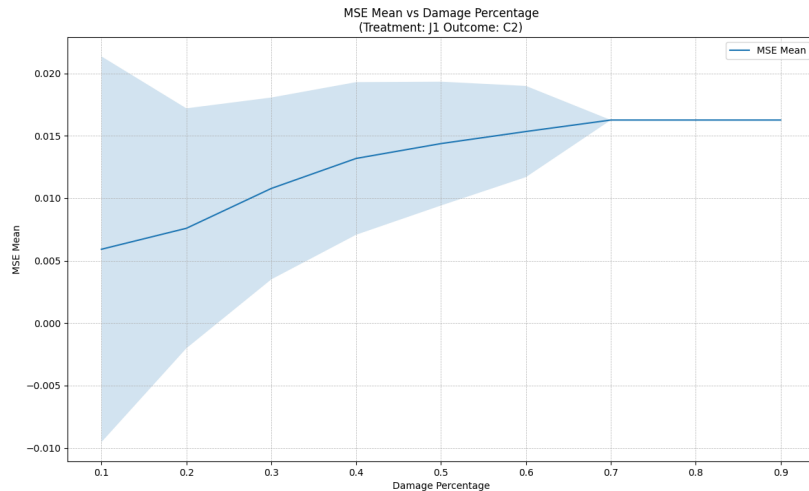


Figure 5.4: Mean Squared Error (MSE) vs damaged percentage with the shaded area indicating the error margin.

The MSE exhibits a distinctive pattern characterised by a sharp initial increase followed by a plateauing effect at higher damage levels. As damage exceeds 60%, the error margins narrow significantly, with MSE values converging towards 0.016 ± 0.04 . This convergence reveals that while predictions become uniformly poor at high damage levels, they also become more consistently inaccurate. Such behaviour likely stems from the simplified network dynamics that emerge when extensive structural compromise reduces the available pathways for water flow.

5.1.4 Implications for Network Management

These findings provide crucial insights into network maintenance and monitoring. The rapid deterioration in prediction accuracy between 10% and 40% damage highlights a critical maintenance window. Network operators should initiate repairs when damage nears 10% to prevent rapid degradation. The identification of a 70% damage threshold as a point of complete prediction failure establishes a clear upper bound for acceptable network damage, though the practical threshold for reliable prediction lies much lower at around 40% damage. The convergence of error margins at high damage levels indicates that prediction reliability becomes uniformly poor, potentially complicating damage assessment in severely compromised networks.

5.2 Research Question 2: Impact of Damage Location on Prediction Accuracy

This section examines how damage location influences prediction accuracy across different network regions. For our analysis, we define mid-network nodes as junction points that serve as intermediate connection points between source nodes (reservoirs)

and end-user consumption points. Specifically, these are nodes J_1 through J_4 in our network, which facilitate water distribution and flow control but do not directly serve as either supply or consumption points.

5.2.1 Regional Vulnerability Analysis

To understand how damage location affects prediction accuracy, Figure 5.5 presents a comparative analysis of ATE across different network regions.

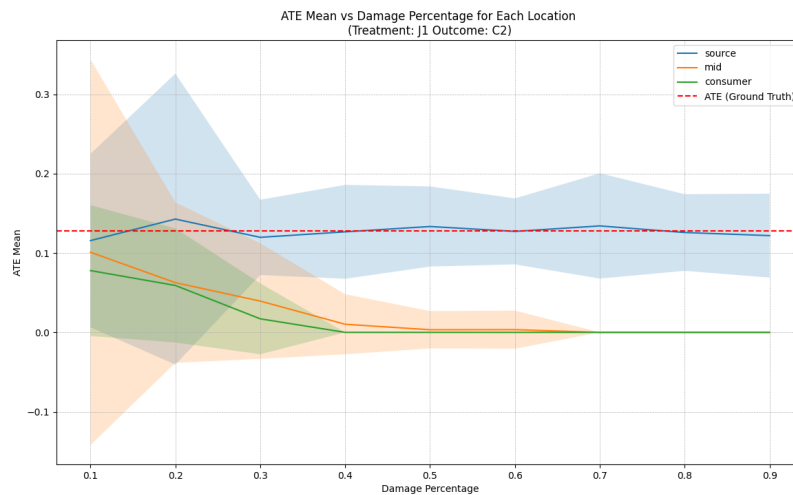


Figure 5.5: A comparison of the ATE vs damage percentages across the different locations (source, junctions/mid and consumer) compared to the ground truth ATE.

The analysis reveals striking differences in network vulnerability based on damage location. Mid-network and consumer nodes demonstrate substantially larger deviations from the ground truth ATE compared to source node damage. Source nodes exhibit greater resilience, showing smaller impacts on ATE, though with notable variability in their response patterns. This variation in vulnerability emphasises the critical role of mid-network junctions as water distribution conduits, where their compromise leads to amplified cascading effects throughout the system. The network’s inherent redundancies at source locations provide some buffering against downstream impacts, suggesting evolutionary pressure in network design has naturally reinforced these critical entry points.

To further quantify these location-specific impacts, Figure 5.6 illustrates the percentage change in ATE across different damage locations.

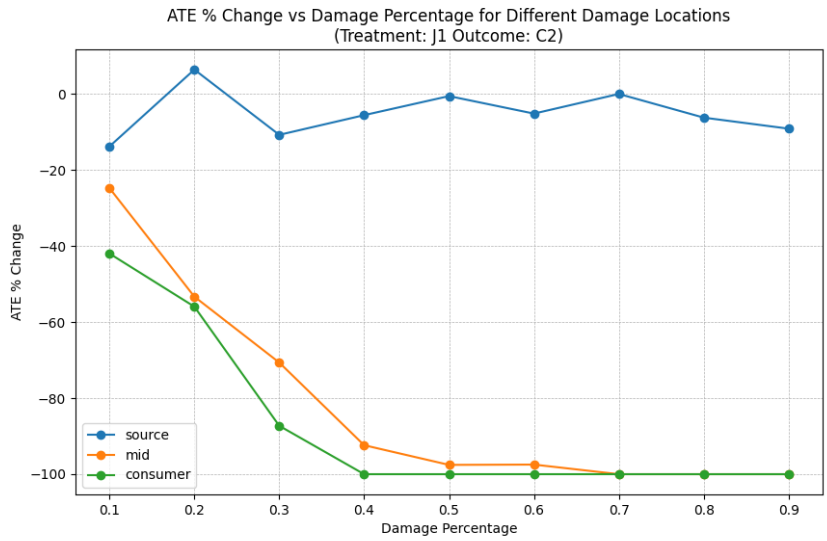


Figure 5.6: Percentage change in ATE vs Percentage damage for the damaged model across different damage locations

The results demonstrate consistently higher percentage changes in ATE for mid-network and consumer node damage compared to source node damage. Mid-network damage produces an average ATE reduction of approximately 45% at just 30% damage, while source nodes maintain relatively stable predictions until reaching higher damage levels. This pattern reinforces the structural significance of these regions in maintaining network integrity and predictive accuracy.

5.2.2 Error Distribution Analysis

Figure 5.7 presents a detailed comparison of MSE across different damage locations, providing additional validation of our vulnerability findings.

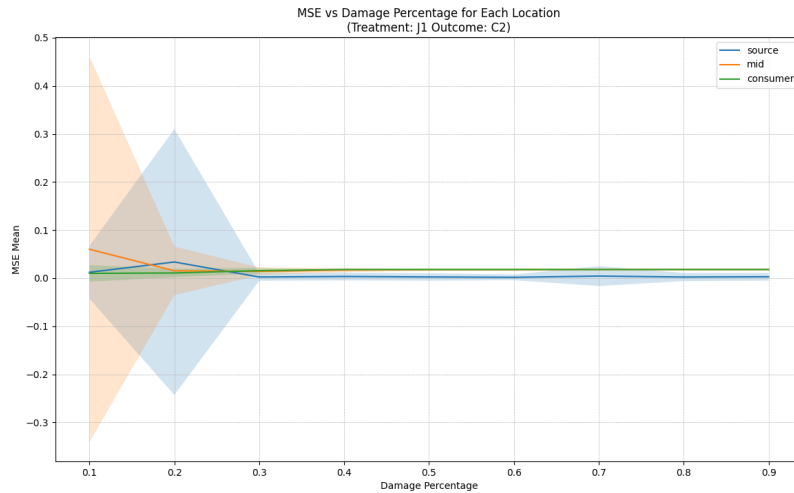


Figure 5.7: A comparison of the MSE vs damage percentages across the different locations (source, junctions/mid and consumer).

The MSE analysis confirms the hierarchical vulnerability pattern, showing markedly higher error values for mid-network and consumer node damage compared to source node damage. Particularly noteworthy are the larger error margins observed at lower damage levels (10-20%) for these critical regions, followed by a stabilisation at uniformly high levels as damage increases. This error pattern suggests that early-stage damage to mid-network and consumer nodes has an outsized impact on prediction reliability.

5.2.3 Treatment Location Effects

The effectiveness of treatment locations in maintaining predictive accuracy reveals additional insights into network dynamics. Figure 5.8 shows ATE measurements at node C_2 against source node damage for different treatment locations.

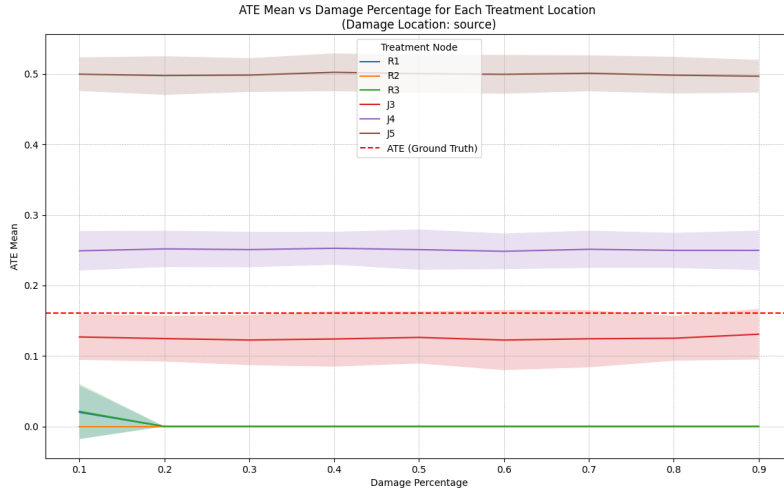


Figure 5.8: A plot of the ATE measured at node C_2 against varying damage to the source nodes for a variety of treatment locations

Mid-network treatment nodes demonstrate behaviour similar to source node damage scenarios, with ATE values maintaining relative stability around specific thresholds. Treatment applied at J_3 yields ATE values closest to the ground truth, while other treatment nodes show significant deviations. Particularly noteworthy is the rapid convergence to zero ATE by 20% damage when treatments are applied to source nodes (R_1 , R_2 , R_3), indicating the complete loss of measurable treatment effects due to pathway disruption. These findings underscore the critical dependence of causal inference accuracy on treatment node location.

Figure 5.9 presents the ATE measurements for mid-network node damage scenarios.

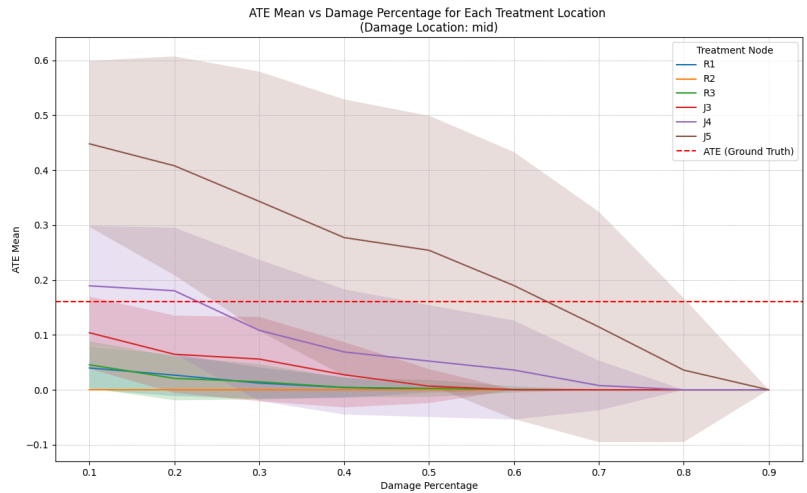


Figure 5.9: A plot of the ATE measured at node C_2 against varying damage to the middle nodes for a variety of treatment locations

Treatment nodes positioned near source points (J_3, J_4) initially demonstrate ATE values closely aligned with ground truth measurements, followed by a gradual deterioration toward zero. This contrasts sharply with source node treatments, which exhibit accelerated convergence to zero ATE due to the rapid reduction in available measurement pathways as mid-network damage increases. This spatial relationship between treatment location and prediction resilience highlights the importance of strategic treatment point placement in maintaining network predictive capability.

The final analysis focuses on consumer node damage scenarios, presented in Figure 5.10.

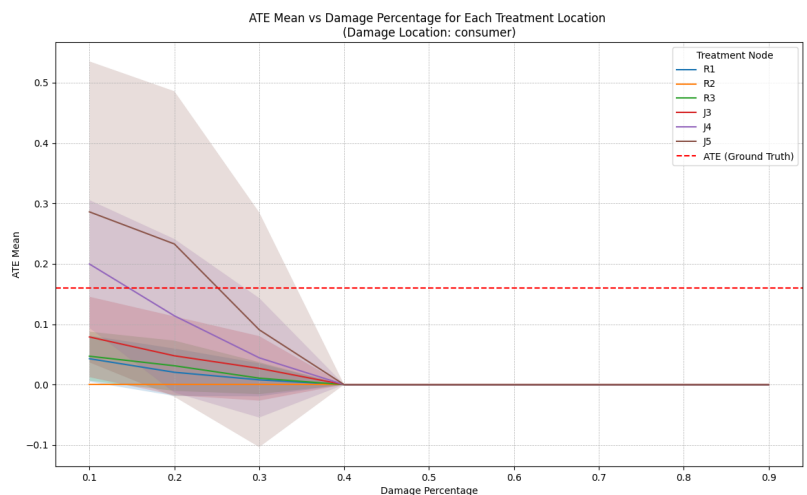


Figure 5.10: A plot of the ATE measured at node C_2 against varying damage to the consumption nodes for a variety of treatment locations

Consumer nodes exhibit the most rapid convergence of ATE to zero under damage scenarios, with predictive capability diminishing almost entirely by 30% damage. This swift decline emphasises their direct dependence on intact upstream pathways and reveals their position as the most vulnerable points in the network for maintaining prediction accuracy.

These location-specific findings reveal that effective network management requires a nuanced understanding of spatial vulnerability patterns. The heightened sensitivity of mid-network and consumer nodes to damage suggests that maintenance strategies should prioritise these regions to maintain system performance. Furthermore, the strong influence of treatment location on prediction accuracy provides valuable guidance for optimising network control points, particularly in scenarios where resources for network monitoring and maintenance are limited.

5.3 Summary of Findings and Research Contributions

This research has advanced our understanding of WDN resilience through systematic investigation of damage impacts on prediction accuracy. Our analysis reveals two fundamental contributions to the field: first, the identification of critical damage thresholds and their non-linear effects on system performance, and second, the characterisation of location-dependent vulnerability patterns that challenge conventional network protection strategies.

The experimental findings demonstrate that network damage most critically impacts prediction accuracy in the 10%-40% range, where the system experiences rapid declines in Average Treatment Effect accompanied by sharp increases in Mean Squared Error. This early-stage sensitivity fundamentally changes our understanding of network vulnerability, suggesting that even minor structural compromises can trigger disproportionate degradation in predictive capability. The convergence of error metrics at higher damage levels (beyond 60%) further indicates a systemic threshold beyond which prediction becomes uniformly unreliable.

Our spatial sensitivity analysis reveals previously unrecognised patterns in network vulnerability. Mid-network and consumer nodes emerge as critical weak points, exhibiting substantially higher prediction errors when damaged compared to source nodes. This finding challenges traditional infrastructure protection paradigms that primarily focus on source point resilience. The observed spatial dependence of intervention effectiveness, where treatments positioned closer to outcome variables maintain better predictive reliability, provides crucial insights for optimising network control strategies.

The methodological framework illustrated in this research extends beyond traditional statistical approaches by explicitly modelling cause-and-effect relationships in network flow dynamics. This innovation enables network operators to quantify structural damage impacts systematically and identify optimal intervention points for network control. The framework's ability to predict cascade effects and characterise location-specific vulnerabilities provides a foundation for evidence-based maintenance prioritisation and resource allocation.

These contributions collectively advance both theoretical understanding of network resilience and practical infrastructure management. By establishing the relationship between structural integrity and predictive capability, while accounting for spatial vulnerability patterns, our research provides a comprehensive basis for developing more resilient water distribution systems. The methodologies developed here show promise for extension to other critical infrastructure networks where understanding causal relationships and predicting system behaviour are essential for reliable operation.

Chapter 6

Conclusion

This research advances the analysis of WDNs through the novel application of established causal inference methodologies. By applying graphical causal models to water infrastructure systems, we demonstrate a successful application of causal inference frameworks to WDN analysis, specifically addressing the gap in quantifying structural damage effects on prediction accuracy. This approach establishes a methodological bridge between causal modelling and infrastructure resilience assessment, advancing beyond the limitations of previous methods identified in Chapter 3.

To contextualise our contribution within existing approaches, Table 6.1 compares our causal inference framework against alternative methodologies discussed in the literature review:

Table 6.1: Comparison of WDN Analysis Methods

Method	Causal Mechanism Representaion	Computational Efficiency	Uncertainty Quantification	Interpretability
Hydraulic Simulation [Tabesh <i>et al.</i> 2009]	Limited	Low	Limited	Moderate
Graph Theory [Yazdani and Jeffrey 2011]	None	High	None	High
Hybrid Approaches [Di Nardo <i>et al.</i> 2018]	Partial	Moderate	Limited	Moderate
Machine Learning [Shen and Osman 2021]	Implicit	Variable	Moderate	Poor
Our Causal Framework	Explicit	Moderate	Good	High

As this comparison illustrates, our causal inference approach specifically addresses the gap in explicit representation of causal mechanisms—a critical limitation in existing methods—whilst maintaining reasonable computational efficiency and strong interpretability. The framework’s ability to quantify uncertainty through probabilistic reasoning provides additional advantages for infrastructure management under conditions of incomplete information.

Our investigation yields three primary contributions to the field. We demonstrate the successful application of causal inference frameworks to WDN analysis, establishing a methodological bridge between causal modelling and infrastructure resilience assessment. This application enables precise quantification of damage effects on prediction accuracy, moving beyond correlational analysis and establishing causal relationships in network behaviour.

Our analysis reveals previously unidentified vulnerability patterns in network structure. Mid-network and consumer nodes exhibit heightened susceptibility to damage, with cascading effects that significantly impair prediction accuracy throughout the system. This finding contrasts with the observed resilience of source nodes, where built-in redundancies provide effective protection against upstream disruptions. The identification of a critical damage threshold (10%-40%) where predictive capability degrades disproportionately provides concrete guidance for maintenance prioritisation.

We establish quantitative relationships between damage location and system resilience, demonstrating how spatial patterns in network damage influence prediction reliability. This spatial characterisation of vulnerability provides actionable insights for infrastructure protection strategies and maintenance planning. Despite these contributions, several limitations warrant consideration. Our application of causal inference assumes stable environmental conditions and perfect observability of network states, conditions rarely met in real-world systems. The current implementation models only complete pipe blockages, whereas actual damage often manifests as partial restrictions to flow. Our analysis focuses solely on steady-state conditions, not accounting for temporal dynamics in water demand and supply patterns.

Additionally, the validation of our findings relies on simulated data, which may not fully capture the complexity of real-world WDNs. The assumption of linear relationships in our causal models, while mathematically tractable, may oversimplify actual network behaviour under stress conditions.

Several promising directions emerge for future research. Extension of causal models to incorporate partial damage scenarios would enhance practical applicability. Integration of temporal dynamics in prediction models could capture fluctuating demand patterns more accurately. Development of robust estimation techniques for partially observable networks would address real-world monitoring limitations. Application of non-linear causal modelling approaches might better represent complex hydraulic behaviours. Validation against real-world networks of varying scales and complexities would strengthen empirical foundations. Integration with real-time monitoring systems could enhance early warning capabilities, while expanding the analysis to include water quality metrics and demand variability would provide a more comprehensive understanding of network dynamics.

The application of causal inference to WDNs demonstrates significant potential for improving infrastructure resilience analysis. This work illustrates the application of a causal framework for understanding how damage propagates through complex networks, equipping network operators with evidence-based strategies for maintenance and reliability enhancement. Our findings particularly emphasize the importance of

protecting mid-network nodes and implementing early detection systems, challenging traditional infrastructure protection paradigms that focus primarily on source points. By identifying critical vulnerabilities and providing a framework for quantifying damage impacts, this research contributes to more resilient water infrastructure systems and offers methodological insights applicable to other networked infrastructure domains.

References

- [Ajoodha 2018] Ritesh Ajoodha. *Influence modelling and learning between dynamic bayesian networks using score-based structure learning*. PhD thesis, 2018.
- [Ali *et al.* 1999] Seema Ali, Jeremy Bell, Chris Budd, Adriana Davidova, Nathan Kricklock, Margaret Liang, Yves Lucet, Scott Maclachlan, Marni Mishna, Shelly Pinder, Peilin Shi, and Cristina Stoica. Contaminant transport in municipal water systems. 11 1999.
- [Berardi *et al.* 2008] Luigi Berardi, Orazio Giustolisi, Zoran Kapelan, and Dragan A. Savic. Development of pipe deterioration models for water distribution systems using epr. *Journal of Hydroinformatics*, 10(2):113–126, 2008.
- [Colombo and Karney 2002] Andrew F. Colombo and Bryan W. Karney. Energy and costs of leaky pipes: Toward comprehensive picture. *Journal of Water Resources Planning and Management*, 128(6):441–450, 2002.
- [Di Nardo *et al.* 2018] Armando Di Nardo, Michele Di Natale, Carlo Giudicianni, Dino Musmarra, Juan Manuel Ramos Varela, Giovanni Francesco Santonastaso, Armando Simone, and Velitchko Tzatchkov. Redundancy features of water distribution systems. *Procedia Engineering*, 186:412–419, 2018.
- [Farahmandfar *et al.* 2020] Zahra Farahmandfar, Kalyan R. Piratla, and Ronald D. Andrus. Resilience evaluation of water supply networks against seismic hazards. *Journal of Pipeline Systems Engineering and Practice*, 11(1):04019046, 2020.
- [Fenton and Neil 2012] Norman Fenton and Martin Neil. *Risk Assessment and Decision Analysis with Bayesian Networks*. CRC Press, Boca Raton, FL, 2012.
- [Gheisi and Naser 2014] A. Gheisi and Gh. Naser. Water distribution system reliability under simultaneous multicomponent failure scenario. *Journal of American Water Works Association*, 106(7):E319–E327, 2014.
- [Gheisi *et al.* 2016] Alireza Gheisi, Mark Forsyth, and Ghansham Naser. Water distribution systems reliability: A review of research literature. *Journal of Water Resources Planning and Management*, 142(11):04016047, 2016.
- [Giustolisi *et al.* 2008] Orazio Giustolisi, Dragan Savic, and Zoran Kapelan. Pressure-driven demand and leakage simulation for water distribution networks. *Journal of Hydraulic Engineering*, 134(5):626–635, 2008.

- [Goodfellow *et al.* 2016] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, Cambridge, MA, 2016.
- [Hu *et al.* 2023] Guangji Hu, Haroon R Mian, Saeed Mohammadiun, Manuel J Rodriguez, Kasun Hewage, and Rehan Sadiq. Appraisal of machine learning techniques for predicting emerging disinfection byproducts in small water distribution networks. *Journal of Hazardous Materials*, 446:130633, 2023.
- [Kabir *et al.* 2015] Golam Kabir, Solomon Tesfamariam, Alex Francisque, and Rehan Sadiq. Evaluating risk of water mains failure using a bayesian belief network model. *European Journal of Operational Research*, 240(1):220–234, 2015.
- [Kepa 2021] Urszula Kepa. Use of the hydraulic model for the operational analysis of the water supply network: A case study. *Water*, 13(3):326, 2021.
- [Koller and Friedman 2009] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, Cambridge, MA, 2009.
- [Li and Li 2021] Yang Li and Xia Li. Research on water distribution systems from the past to the future: a bibliometric review. *Environmental Technology Reviews*, 10(1):161–176, 2021.
- [Ma *et al.* 2024] Haixiang Ma, Xuechun Wang, and Dongsheng Wang. Pump scheduling optimization in urban water supply stations: A physics-informed multiagent deep reinforcement learning approach. *International Journal of Energy Research*, 2024(1):9557596, 2024.
- [Mays 2000] Larry W. Mays. *Water Distribution Systems Handbook*. McGraw-Hill, New York, 2000.
- [Morosini *et al.* 2020] Attilio Francesco Morosini, Francesco Costanzo, Paolo Veltri, and Dragan Savic. Identification of hydraulic criticalities in water distribution systems using graph-theoretic algorithms. *Water*, 12(1):168, 2020.
- [Mounce *et al.* 2011] S. R. Mounce, R. B. Mounce, and J. B. Boxall. Novelty detection for time series data analysis in water distribution systems using support vector machines. *Journal of Hydroinformatics*, 13(4):672–686, 2011.
- [Nault and Karney 2020] Josée Nault and Bryan Karney. Improved rigid water column formulation for simulating slow transients and controlled operations. *Journal of Hydraulic Engineering*, 146(2):04019066, 2020.
- [Newman 2010] Mark Newman. *Networks: An Introduction*. Oxford University Press, Oxford, UK, 2010.
- [Ostfeld *et al.* 2008] Avi Ostfeld, James G. Uber, Elad Salomons, Jonathan W. Berry, William E. Hart, Cindy A. Phillips, Jean-Paul Watson, Gianluca Dorini, Philip Jonkergouw, Zoran Kapelan, Francesca di Pierro, Soon-Thiam Khu, Dragan Savic, Demetrios Eliades, Marios Polycarpou, Suresh R. Ghimire, Brian D. Barkdoll, Roberto Gueli, Jinhui J. Huang, Edward A. McBean, William James, Andreas Krause, Jure Leskovec, Shannon Isovitsch, Jianhua Xu, Carlos Guestrin, Jeanne VanBriesen, Mitchell Small, Paul Fischbeck, Ami Preis, Marco Propato, Olivier

- Piller, Gary B. Trachtman, Zheng Yi Wu, and Tom Walski. The battle of the water sensor networks (bwsn): A design challenge for engineers and algorithms. *Journal of Water Resources Planning and Management*, 134(6):556–568, 2008.
- [Pearl and Verma 1995] Judea Pearl and Thomas S Verma. A theory of inferred causation. In *Studies in Logic and the Foundations of Mathematics*, volume 134, pages 789–811. Elsevier, 1995.
- [Pearl 1988] Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan kaufmann, 1988.
- [Pearl 2009a] Judea Pearl. Causal inference in statistics: An overview. 2009.
- [Pearl 2009b] Judea Pearl. *Causality*. Cambridge university press, 2009.
- [Pearl 2009c] Judea Pearl. *Causality*. Cambridge University Press, Cambridge, UK, 2009.
- [Pearl 2014] Judea Pearl. Graphical models for probabilistic and causal reasoning. *Computing Handbook, 3rd ed.(1)*, pages 44–1, 2014.
- [Phan *et al.* 2019] Thuc D Phan, James CR Smart, Ben Stewart-Koster, Oz Sahin, Wade L Hadwen, Lien T Dinh, Iman Tahmasbian, and Samantha J Capon. Applications of bayesian networks as decision support tools for water resource management under climate change and socio-economic stressors: a critical appraisal. *Water*, 11(12):2642, 2019.
- [Rajani and Kleiner 2001] Balvant Rajani and Yehuda Kleiner. Comprehensive review of structural deterioration of water mains: physically based models. *Urban Water*, 3(3):151–164, 2001.
- [Rossman 2000] Lewis A. Rossman. Epanet 2: users manual. *US Environmental Protection Agency. Office of Research and Development. National Risk Management Research Laboratory*, 2000.
- [Sanz *et al.* 2016] Gerard Sanz, Ramon Pérez, Zoran Kapelan, and Dragan Savic. Leak detection and localization through demand components calibration. *Journal of Water Resources Planning and Management*, 142(2):04015057, 2016.
- [Shen and Osman 2021] Hao Shen and Hesham Osman. Deep learning for predicting the occurrence and location of pipe bursts in water distribution systems. *Journal of Water Resources Planning and Management*, 147(6):04021031, 2021.
- [Shuang *et al.* 2019] Qing Shuang, Mingyuan Zhang, and Yixing Yuan. Performance and reliability analysis of water distribution systems under cascading failures and the identification of crucial pipes. *PLoS ONE*, 14(4):e0215868, 2019.
- [Spirtes *et al.* 2000] Peter Spirtes, Clark N. Glymour, and Richard Scheines. *Causation, Prediction, and Search*. MIT Press, Cambridge, MA, 2000.
- [Tabesh *et al.* 2009] M. Tabesh, A. Hadi Abdolghafar Yekta, and R. Burrows. An integrated model to evaluate losses in water distribution systems. *Water Resources Management*, 23:477–492, 2009.

- [The Goan Network] The Goan Network. Damage to pipeline disrupts water supply in dharbandora. *The Goan Network*.
- [Todini and Pilati 1988] Ezio Todini and Stefano Pilati. A gradient algorithm for the analysis of pipe networks. pages 1–20, 1988.
- [Truong *et al.* 2024] Huy Truong, Andrés Tello, Alexander Lazovik, and Victoria Degeler. Graph neural networks for pressure estimation in water distribution systems. *Water Resources Research*, 60(7):e2023WR036741, 2024.
- [Walski *et al.* 2003] Thomas M. Walski, Donald V. Chase, Dragan A. Savic, Walter Grayman, Stephen Beckwith, and Edmundo Koelle. *Advanced Water Distribution Modeling and Management*. Haestad Press, Waterbury, CT, 2003.
- [Yazdani and Jeffrey 2011] A. Yazdani and P. Jeffrey. Complex network analysis of water distribution systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21(1):016111, 2011.
- [Zanfei *et al.* 2022] Ariele Zanfei, Bruno M Brentan, Andrea Menapace, Maurizio Righetti, and Manuel Herrera. Graph convolutional recurrent neural networks for water demand forecasting. *Water Resources Research*, 58(7):e2022WR032299, 2022.
- [Zeng *et al.* 2017] F. Zeng, X. Li, and K. Li. Modeling complexity in engineered infrastructure system: Water distribution network as an example. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(2):023105, 2017.