Pipe failure prediction and impacts assessment in a water distribution network

Submitted by Konstantinos Kakoudakis to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Engineering in March 2019

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: ……  ……..

Abstract

Water distribution networks (WDNs) aim to provide water with desirable quantity, quality and pressure to the consumers. However, in case of pipe failure, which is the cumulative effect of physical, operational and weather-related factors, the WDN might fail to meet these objectives. Rehabilitation and replacement of some components of WDNs, such as pipes, is a common practice to improve the condition of the network to provide an acceptable level of service.

The overall aim of this thesis is to predict—long-term, annually and short-term—the pipe failure propensity and assess the impacts of a single pipe failure on the level of service. The long-term and annual predictions facilitate the need for effective capital investment, whereas the short-term predictions have an operational use, enabling the water utilities to adjust the daily allocation and planning of resources to accommodate possible increase in pipe failure. The proposed methodology was implemented to the cast iron (CI) pipes in a UK WDN.

The long-term and annual predictions are made using a novel combination of Evolutionary Polynomial Regression (EPR) and K-means clustering. The inclusion of K-means improves the predictions' accuracy by using a set of models instead of a single model. The long-term predictive models consider physical factors, while the annual predictions also include weather-related factors. The analysis is conducted on a group level assuming that pipes with similar properties have similar breakage patterns. Soil type is another *aggregation* criterion since soil properties are associated with the corrosion of metallic pipes.

The short-term predictions are based on a novel Artificial Neural Network (ANN) model that predicts the variations above a predefined threshold in the number of failures in the following days. The ANN model uses only existing weather data to make predictions reducing their uncertainty.

The cross-validation technique is used to derive an accurate estimate of accuracy of EPR and ANN models by guaranteeing that all observations are used for both training and testing, and each observation is used for testing only once.

The impact of pipe failure is assessed considering its duration, the topology of the network, the geographic location of the failed pipe and the time. The performance indicators used are the ratio of unsupplied demand and the number of customers with partial or no supply. Two scenarios are examined assuming that the failure occurs when there is a peak in either pressure or demand. The pressure-deficient conditions are simulated by introducing a sequence of artificial elements to all the demand nodes with pressure less than the required.

This thesis proposes a new combination of a group-based method for deriving the failure rate and an individual-pipe method for evaluating the impacts on the level of service. Their conjunction indicates the most critical pipes.

The long-term approach improves the accuracy of predictions, particularly for the groups with very low or very high failure frequency, considering diameter, age and length. The annual predictions accurately predict the fluctuation of failure frequency and its peak during the examined period. The EPR models indicate a strong direct relationship between low temperatures and failure frequency.

The short-term predictions interpret the intra-year variation of failure frequency, with most failures occurring during the coldest months. The exhaustive trials led to the conclusion that the use of four consecutive days as input and the following two days as output results in the highest accuracy. The analysis of the relative significance of each input variable indicates that the variables that capture the intensity of low temperatures are the most influential.

The outputs of the impact assessment indicate that the failure of most of the pipes in both scenarios (i.e. peak in pressure and demand) would have low impacts

(i.e. low ratio of unsupplied demand and small number of affected nodes). This can be explained by the fact that the examined network is a large real-life network, and a single failure of a distribution pipe is likely to cause pressure-deficient conditions in a small part of it, whereas performance elsewhere is mostly satisfactory. Furthermore, the complex structure of the WDN allows them to recover from local pipe failures, exploiting the topological redundancy provided by closed loops, so that the flow could reach a given demand node through alternative paths.

Contents

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors, Associate Professor Raziyeh Farmani and Professor David Butler, whose guidance, expertise and encouragement throughout my studies have been invaluable. I would also like to thank all the staff and students in the Centre for Water Systems, both past and present.

I would like to thank all my friends and family who have encouraged and supported me for the last four years.

List of Figures

List of Tables

List of Abbreviations

AC	Asbestos Cement
ANN	Artificial Neural Network
AUC	Area under Curve
BBN	Bayesian Belief Network
BMAM	Bayesian Model Averaging Method
BNHPP	Bayesian Non-Homogenous Poisson Process
BWPPM	Bayesian Weibull Proportional Hazard Model
BZNHPP	Bayesian zero-inflated Non-Homogenous Poisson Process
CI	Cast Iron
CoD	Coefficient of Determination
CML	Customers Minutes Lost
CP	Cathodic protection
CWWA	Canadian Water and Wastewater Association
DDA	Demand-driven analysis
DI	Ductile Iron
ERM	Exponential regression model
EPR	Evolutionary Polynomial Regression
EPS	Extended Simulation Period
FPR	False Positive Rate

FI	Freezing Index
GIS	Geographical Information System
GLMs	Generalized linear models
GP	Genetic Programming
GSS	Guaranteed Standard Scheme
HAD	Head-driven analysis
HFR	Head-outflow relationship
KPIs	Key Performance Indicators
LEYP	Linear Extension Yule Process
LGLM	Logistic Generalized Linear Regression Model
LGM	Linear regression model
LS	Least Square method
HFES	Hierarchical Fuzzy Expert System
MCDA	Multi- Criteria Decision Analysis
MERM	Multivariate Exponential Regression Model
MLRM	Multivariate Linear Regression Model
MOGA	Multi-Objective Genetic Algorithm
M/R/R	Maintenance/Rehabilitation/Replacement
NHFR	Nodal Head-Flow Relationship
NHPP	Non-Homogenous Poisson Process

NOPF	Number of Previous Failures
OFWAT	Office of Water Services
PDA	Pressure-driven analysis
PE	Polyethylene
PGLRM	Poisson Generalized Linear Regression Model
PHM	Proportional Hazard Model
PLP	Power Law Process
PM	Poisson process Model
PVC	Polyvinyl chloride
ROC	Receiver Operating Curve
ROCOF	Rate of Occurrence of Failure
RMSE	Root Mean Square Error
R^2	Coefficient of Determination
TPR	True Positive Rate
WALM	Weibull Accelerated Lifetime Model
WDN	Water Distribution Network
WDS	Water Distribution System
WPHM	Weibull Proportional Hazard Model
ZINHPP	Zero-Inflated Non-Homogenous Poisson Process

Chapter 1. Introduction and Scope

1.1 Background

Water distribution networks are critical infrastructures (Meng et al. 2018) that aim to provide water with desirable quantity, quality and pressure to the consumers (Tscheikner-Gratl et al. 2017) —even under abnormal functioning conditions (e.g. pipe failure) (Xu and Goulter 1999).

The failure of water pipes is the result of structural weakening coupled with externally and internally imposed stresses (Sadiq et al. 2004b) and leads to environmental, economic and social costs (Kunkel et al. 2008; Haider et al. 2013). An average of 850 water main failures occur daily in North America, with a total annual repair cost of more than $3 billion (Baird 2011). A United States Environmental Protection Agency (USEPA) survey estimated the cost requirements for upgrading water distribution and transmission systems in the United States at US $77 billion for a 20 years period (Davies et al. 1997). In Canada, the Canadian Water and Wastewater Association estimated that CAN $11.5 billion would be required for water main upgrading over the next 15 years (CWWA 1997). It is estimated that the annual maintenance cost for the water distribution network in Australia exceeds A$1.4 billion (in 2012) for a network of approximately 163,000 km with a total asset value exceeding A$71.1 billion dollars (BITRE 2014). Typically, water pipes are being rehabilitated at an annual rate of 0.5±1% of the existing length of the distribution system to prevent further ageing.

In the UK, OFWAT monitors the standard of delivered potable water in UK using a number of key performance indicators (KPIs), focusing on long-term pressure adequacy (i.e. pressure of water mains, the DG2 indicator) and continuity of water

supply (i.e. supply interruptions, the DG3 indicator). OFWAT (2008) describes the DG2 indicator as: “The register must clearly identify those properties reported under DG2 and distinguish them from those that receive low pressure but are excluded from DG2 and provide a verifiable reason for the exclusion (e.g. as abnormal demand or short duration of low pressure).” The DG3 indicator is described as (OFWAT 2008): “The aim of this indicator is to identify the number of properties affected by planned and unplanned supply interruptions lasting longer than 3 hours, 6 hours, 12 hours and 24 hours.”

In some cases, a WDN may fail to meet the objectives, and the failures can be categorized into the following types (Ostfeld et al. 2002; Ozger 2003): 1) performance failure, 2) mechanical failure and 3) water quality failure. The most common type of mechanical failure is pipe failure (Ozger 2003).

An effective asset management plan enables water utilities to optimize investment in their assets through better strategic and capital planning processes (Zamenian et al. 2016; D’Ercole et al. 2018). However, the water pipes are usually buried underground, and, therefore, the direct monitoring and inspection for obtaining adequate data for use as a basis for deterioration forecasting analysis is difficult (Rogers and Grigg 2009; Liserra et al. 2014; Shin et al. 2016; Tscheikner-Gratl et al. 2016; Salehi et al. 2018). Because of these difficulties, predictive models have been developed to predict the likelihood of pipe failure proactively and assist in the asset management plans (Lim et al. 2008; Herstein et al. 2010). The pipe failure can have an impact (among others) on the level of service (Giustolisi et al. 2016). The magnitudes and the scales of the impacts depend on many factors, amongst which, geographic location of pipe failure, the time of pipe failure and its duration and the topology and complexity of the WDN are some of the most important (Bicik 2010).

1.2 Aims and Objectives

The overall aim of the thesis is to predict the pipe failure long-term, annually and short-term and then assess the impacts due to a single pipe failure on the level of service. The long-term and annual predictions facilitate the need for decision-making to enhance the condition of water pipes to deliver an acceptable level of service and capital investment. The short-term predictions have an operational use for planning of resources to accommodate possible increases in pipe failure. The overall aim of the thesis will be achieved through the following objectives:

1. To accurately capture the failure patterns in a WDN, combining a data-driven and a clustering method
2. To examine the annual variation of pipe failure frequency considering both weather-related and physical factors
3. To derive the failure rate of individual pipes through the *aggregated* predictive models
4. To associate the failure propensity with pipe characteristics and the soil type
5. To visualize the most prone to failure parts of the network and make the outputs of the predictive models more understandable
6. To predict the occurrence of a large number of failures in the short-term considering weather-related factors
7. To simulate the pressure-deficient conditions (pressure-driven analysis) by introducing a series of artificial elements to all the demand nodes with pressure less than the required
8. To estimate the impacts caused to the level of service due to a single pipe failure using the ratio of unsupplied demand and the number of nodes with zero and partial supply

1.3 Thesis Structure

Chapter 1 Introduction and Scope

Background information is presented, and the overall aim and the objectives of the research are detailed. The originality of the thesis and the contribution to knowledge are highlighted.

Chapter 2 Literature Review

The literature review presents and categorizes the factors that contribute to pipe failure. It provides an in-depth analysis of the models for predicting the pipe failure and calculating the impacts. The predictive models are analysed with respect to their suitability for accurately capturing the failure patterns and the impact assessment models with respect to their ability to quantify severity of the impacts. The limitations of the existing methods are discussed, and, finally, the advantages of the proposed methodology are highlighted.

Chapter 3 Pipe Failure Prediction

Three models for predicting pipe failure long-term, annually and short-term are presented. The long-term and the annual predictions rely on a novel combination of EPR and K-means clustering method. The long-term predictions consider physical factors as explanatory variables while the annual predictions are novel in considering both weather-related and physical factors. The EPR models are used to derive the failure rate of individual pipes and identify those most prone to failure. The short-term predictions are based on a novel ANN model that predicts the following days with a large number of failures using existing weather data as input. Various combinations of inputs and outputs are examined to select the one with the highest accuracy.

This chapter is based on the following publications:

Kakoudakis, K., R. Farmani and D. Butler (2018). Pipeline failure prediction in water distribution networks using weather conditions as explanatory factors. Journal of Hydroinformatics, 20 (5), 1191-1200

Kakoudakis, K., K. Behzadian, R. Farmani and D. Butler (2017). Pipeline failure prediction in water distribution networks using evolutionary polynomial regression combined with K-means clustering. Urban Water Journal, 14 (7), 737-742

Farmani, R., K. Kakoudakis, K. Behzadian and D. Butler (2017). Pipe Failure Prediction in Water Distribution Systems Considering Static and Dynamic Factors. Procedia Engineering, 186, 117-126

Chapter 4 Impacts Assessment

This chapter entails a section for the pressure-driven analysis. Following this, a method proposed by Mahmoud et al. (2017) for simulating the pressure-deficient conditions caused by pipe failure using a sequence of artificial elements is proposed. The satisfied demand is linked to available pressure. The magnitude of the impacts on the level of service is assessed considering factors such as the geographic location of pipe failure, the time of occurrence and its duration. The performance indicators employed are the fraction of unsupplied demand, the number of nodes with partial supply and the number of nodes with zero supply.

Chapter 5 Case Study

This chapter entails the main characteristics of the case study. This is followed by the pre-processing of the available data and the outputs of the preliminary analysis for pipe *aggregation*. The proposed methodology was implemented to the cast iron (CI) pipes of a part of a WDN in the UK.

Chapter 6 Results of Predictive Models

This chapter provides the results of the implemented methodology for pipe failure prediction. The developed predictive models provide an insight into the relationship between pipe failure and the factors that contribute to it, and their outputs are employed to derive the failure rate of individual pipes. The selection of the explanatory variables is examined in conjunction with the engineering knowledge. The combination of a data-driven and clustering methods resulted in higher accuracy compared to the single-model approach, as the high accuracy enabled the derivation of the failure rate of individual pipes. The annual predictions precisely predict the failure variation and can identify potential peaks. Short-term predictions use recorded data (i.e. do not have to be forecasted) as input, reducing significantly the uncertainty in the predictions.

Chapter 7 Results of Impacts Assessment

The outputs of the impacts assessment indicate that most of the pipes in both scenarios result in low impacts. This can be explained by the fact the examined network is a large real-life network and a single failure of a distribution pipe is likely to cause pressure-deficient conditions in a small part of it, whereas performance elsewhere in the network is mostly satisfactory. Also, the redundant design enables the system to overcome local pipe failures by using alternative paths for supplying demand nodes.

Chapter 8 Discussion, Conclusions and Recommendations

This chapter draws upon the previous ones to present the key research findings and offers recommendations for future research. The conclusions derived, the key findings and the novelty existing in the thesis are further discussed.

1.4 Originality and Contribution to Knowledge

The main contributions and originality of the work presented in this thesis are as follows:

- A new approach has been developed by combining EPR and K-means clustering method to predict the failure patterns in the WDN
- For the first time, the K-means method has been used for creating clusters considering pipe attributes as criteria
- *Aggregated* predictive models have been used to derive the failure rate of individual pipes, which is associated with their characteristics and the soil type
- The annual number of failures has been predicted using a data-driven and a clustering method, considering, for the first time, both weather-related and physical factors
- The Jenks Natural Breaks method and the ArcGIS tool have been used to visualize the outputs of the predictive models and make them more understandable
- Short-term failure predictions have been made to alert the water authorities of an increased number of pipe failures, considering, for the first-time, weather-related factors that do not have to be predicted
- The pressure-deficient conditions are simulated using a sequence of artificial elements based on the approach proposed by Mahmoud et al. (2017). This thesis proposes a new combination of a grouped-based method for deriving the failure rate and an individual-pipe method for evaluating the impacts on the level of service.

1.5 Key Assumptions

The key assumptions made in this thesis are:

1. The aspects of water quality (e.g. discolouration) are not considered, as this thesis is only concerned with the performance assessment with respect to quantity

2. Pipe failure implies a cost for rehabilitation/replacement and can possibly cause loss of business, costs associated with emergency response and damage to other existing nearby infrastructures. Those impacts are not considered, either.

3. It is assumed that there is no intervention (e.g. closure of valves) for 24 hours after the pipe failure occurrence

4. The impacts on the level of service are evaluated based only on the available nodal flow, and the type and sensitivity of customers fed by each node are ignored since this knowledge is not available

Chapter 2. Literature Review

2.1 Introduction

The literature review describes the developed approaches for pipe failure prediction and their impact assessment. In the WDN's case, pipe failure may affect its ability to maintain the pressure within specific limits and satisfy the customer demand. The literature review chapter is composed of four parts. The first part describes the factors that contribute to the deterioration and failure of water pipes. The understanding of those factors aids in pipe failure prediction. The second part is on the methods for predicting the pipe failure, describing their benefits and their shortcomings. The third part entails the developed approaches for calculating the impacts on the level of service due to pipe failure. Finally, are summarised the key points and the gaps of the existing literature. The last part explains how the proposed methodology is aiming to bridge those identified gaps.

2.2 Factors affecting pipe failure

Pipe failure is the cumulative effect of several factors (Table 2.1) acting on the them (Clark et al. 2010). These factors can be classified into three categories (Kleiner and Rajani 2002; Demissie et al. 2017): physical, environmental and operational. Environmental and physical factors can be further divided into static (non-time-dependent) and dynamic (time-dependent), while the operational factors are inherently dynamic. Most of the factors related to the pipe properties (e.g. material, diameter, length) tend to be static, whereas the environmental factors can be either random or cyclical over time (Kleiner and Rajani 2002). Soil

can't be explicitly characterised as a static or dynamic factor since some of the properties (e.g. moisture, temperature) vary over time.

Table 2.1 Factors contributing to pipe failure (Al-Barqawi and Zayed 2006)

Physical	**Environmental**	**Operational**
Pipe material	Climate	Water pressure
Pipe diameter	Climate Change	Previous failures
Pipe age	Soil	Corrosion
Pipe length	Groundwater level	Cathodic Protection
Pipe wall thickness	External loads	Water quality
	Pipe bedding	

2.2.1 Physical Factors

Material

The majority of WDNs consist of cast iron (CI), ductile iron (DI), asbestos cement (AC), polyvinyl chloride (PVC), polyethylene (PE) and concrete pipes. Due to the material properties, pipes made of different materials are expected to deteriorate and fail in different ways (Greyvenstein and Van 2007). Corrosion has been identified as a main cause of failure of metallic (CI and DI) pipes (Makar 2000; Rajani and Kleiner 2001; Spickelmire 2002; Li and Mahmoodian 2013; Ji et al. 2017). AC pipes are vulnerable to chemical reaction by certain soils and aggressive water such as low PH water (Kleiner and Rajani 2001; Reed et al. 2007; Davis et al. 2008). PVC and PE pipes are susceptible to permeation or degradation by certain organic contaminants (Davis et al. 2007; Clair and Sinha 2014). Concrete pipes are vulnerable to chemical attacks from certain aggressive soils and waters (Reed et al. 2007).

Diameter

Pipes with small diameter are expected to have an above the average failure rate (Clark et al. 1982; Walski and Pellicia 1982; Rajani and Tesfamariam 2004; Boxall et al. 2007). Kettler and Goulter (1985) associated the higher failure frequency of small diameter pipes with the reduced pipe strength and wall thickness, the different construction standards and the less reliable joints.

Age

Age is considered as one the major factors that contribute to pipe deterioration (Berardi et al. 2008; Li et al. 2014) and was the first factor to be considered for pipe failure prediction (e.g. Shamir and Howard 1979; Clark et al. 1982). Age gives an indication of the duration a pipe has been laid and it is exposed to external loads and the surrounding environment conditions (Boxall et al. 2007). However, age on its own is a poor indicator and does not have a decisive influence on the optimal point of pipe rehabilitation (Boxall et al. 2007; Malm et al. 2012).

Length

Pipe length is a surrogate for exposure to external and internal stresses and higher exposure is expected to lead to more breaks (Boxall et al. 2007; Yamijala et al. 2009), although the inherent randomness in the relationship between length and breaks is relatively high (Kleiner and Rajani 2010). The total pipe length has been used as a normalizing factor in the group level analysis (e.g. Shamir and Howard 1979; Walski and Pelliccia 1982; Kettler and Goulter 1985; Kleiner and Rajani 2004) implying that the breaks are distributed uniformly and proportionally along the length of the pipes (Kleiner and Rajani 2012).

Wall Thickness

The ability of a pipe to resist forces is a function of the material strength and its geometrical proportions (Skipworth et al. 2002). Thinner walls are affected more from high level stresses for the same external loads (Røstum 2000).

2.2.2 Environmental Factors

Climate

Previous studies have analysed the effects of climate on pipe failure trends in Canada and northern USA (Rajani et al. 1996; Kleiner and Rajani 2002; Rajani and Tesfamariam 2004; Hu and Hubble 2007; Rajani et al. 2012; Laucelli et al. 2014), southern USA (Hudak et al. 1998), Australia (Constantine et al. 1996; Gould et al. 2011b), United Kingdom (Newport 1981; Boxall et al 2007), Netherlands (Wols and Thienen 2013, 2014; Wols et al. 2014) and Austria (Fuchs-Hanusch et al. 2013).

Newport (1981), Fuchs-Hanusch et al. (2013), Wols and Thienen (2013), observed a high correlation between the severity of winter and the pipe failure frequency. Rajani et al. (1996; 2012) associated the increased number of pipe breaks during late fall and early spring with the temperature difference between the water and the soil-backfill close to the pipe and the air temperature transits from above 0^0C to below 0^0C or vice versa. Rajani and Tesfamariam (2004) also observed a marked increase in pipe failure frequency in the presence of high temperature differences between the water in the pipe (1-2^{0}C) and adjacent soil (10-12^{0}C). Hudak et al. (1998) first observed that the failure peak during summer followed extreme dry periods. Gould et al. (2011b) attributed the summer failures peak to a peak in circumferential failures due to the differential soil movement whereas no remarkable increase on longitudinal failures was observed.

Climate Change

Climate change may have an impact on pipe breakage rate since it implies the modification of ‘’average’’ climatic conditions and a more frequent occurrence of extreme events. The effects of climate change have not been documented yet.

Groundwater Level

The lowering of groundwater level and the consolidation of the soil following extended dry periods can result in increased differential settlements that may damage the buried pipes (Wols and Thienen 2014).

Soil

Gould et al. (2011b) linked the peak of circumferential failures with the differential soil movement resulting from soil shrinkage occurring as soil moisture decreases. In general, soils close to the ground surface are prone to atmospheric forcing that leads to greater variation in moisture content and temperature (Gould et al. 2011a; Rajeev et al. 2012). A stronger correlation has been observed in more expansive soils (Hudak et al. 1998; Hu and Hubble 2007; Gould et al. 2011b).

Pipe Bedding

Pipe bedding acts as a surrogate for increased external loading during construction practices that could increase the break rate (Jenkins et al. 2014). Goulter and Kazemi (1988) attributed the observed significant temporal and spatial clustering of water main failures to deteriorated bedding conditions around the failure locations due to leakage of water. The differential soil movement caused by the leaking water results in lack of continuous support beneath the pipes creating the bedding stresses.

External Loads

The external loads applied to the pipes include the traffic loads. Traffic loads develop non-uniform stress conditions around the pipe circumference diminishing the uniform support over its entire length (Rajeev et al. 2014).

2.2.3 Operational Factors

Water Pressure

Excessive pressure above the normal operational conditions or sudden pressure transition can cause failure to the WDN (Skipworth et al. 2002; Greyvenstein and Van 2007). The maximum hydraulic pressure (e.g. Asnaashari et al. 2009 and Ghorbanian et al. 2016), the average hydraulic pressure (e.g. Tabesh et al. 2009), or combination of different values of the hydraulic pressure (e.g. Shirzad et al. 2014; Martínez-Codina et al. 2015a) have been examined in the pipe failure analyses.

Corrosion

External corrosion has been identified as the main deterioration mechanism on the exterior of metallic (CI and DI) pipes and has been associated with the properties of the surrounding soil (Rajani and Kleiner 2003; Sadiq et al. 2004a; Liu et al. 2009; Ji et al. 2017). Gould et al. (2011b) has linked the rate of corrosion with soil corrosivity which increases as soil resistivity decreases (Sadiq et al. 2004a) and soil resistivity decreases as soil moisture content increases (Zhou et al. 2001). Corrosion can be accelerated by high temperatures (Rajani and Kleiner 2001) or by a high temperature variation over a short period (McNeill and Edwards 2002).

Previous Failures

The number of previous failures has been identified as a significant factor in failure prediction in the literature (Clark et al. 1982; Andreou et al. 1987; Le Gat and Eisenbeis 2000; Røstum 2000; Park and Longanathan 2002; Asnaashari et al. 2009). Previous failure(s) may affect pipe failure due to soil movement caused by the changing moisture content from leaking water, or exposure of the soil to the extreme cold air and disturbance of the bedding during rehabilitation (Skipworth et al. 2002).

Cathodic Protection

Cathodic protection (CP) is defined as the reduction or elimination of corrosion and is regarded as an effective method to reduce breakage frequency and extend the useful life of pipes (Kleiner et al. 2003). CP can be applied in two forms, namely, hot spot and retrofit. Hotspot CP is the practice of installing a protective anode at the location of a pipe failure, right after the failure has been repaired, while Retrofit CP refers to the practice of systematically protecting existing pipes with sacrificial anodes (Kleiner and Rajani 2004; Rajani and Kleiner 2007). Schuster and McBean (2008) observed that the probability of a pipe break occurring 10 years after the application of CP decreases in comparison with pipes that have not been cathodically protected.

Water Quality

The chemical quality of the water may attack the AC pipes and cause either reduction in thickness or loss of strength (Hu and Hubble 2007). The water quality can also affect the rate of internal corrosion of metallic pipes (Sander et al. 1996; Yamini and Lence 2010).

2.3 Predictive Models

Pipe failure implies a decrease in the service level, resulting in economic, environmental and social costs (Giustolisi et al. 2006; Li et al. 2015; Martínez-Codina et al. 2015b). In order to cope with the impacts of pipes failures water utilities usually follow one of the two rehabilitation strategies: reactive or proactive (Røstum 2000). In the reactive strategy, a pipe will be rehabilitated after failure is detected, whereas in the proactive strategy pipe rehabilitation is scheduled in advance after assessing and forecasting pipe propensity to fail. The reactive approach is less efficient (Carrión et al. 2010; Debón et al. 2010) and therefore researchers and practitioners have strived to develop predictive models in which the likelihood of failure in pipes is predicted in advance for effective Maintenance/Rehabilitation/Replacement (M/R/R) plans (Dandy and Engelhardt 2006; Dridi et al. 2009; Nafi and Kleiner 2010; Alvisi and Franchini 2010).

The predictive models can be classified as physical (Rajani and Kleiner 2001; Wilson et al. 2017), statistical (Kleiner and Rajani 2001; Nishiyama and Filion 2013; Scheidegger et al. 2015) and data-driven methods which include ANN (Clair and Sinha 2012; Nishiyama and Filion 2013), Genetic Programming (GP) (Xu et al. 2011a; 2011b) and EPR (Giustolisi and Savic 2006; Berardi et al. 2008).

A preliminary distinction must be made between physically based and the rest of the approaches. Physical models analyse the loads to which the pipes are subject to and the capacity of the pipes to resist these loads to predict their propensity to break (Rajani and Kleiner 2001; Tesfamariam et al. 2006). In spite of having a reasonable accuracy, physical models, compared to other methods, have significant input data demands (Wood and Lence 2009) due to the fact that they

try to simulate the mechanisms that lead to pipe failure. These demands involve gaining an understanding of structural behaviour of buried pipes, pipe-soil interaction and knowledge about the quality of installation, internal and external stresses and material deterioration. Physical-based models are also time consuming to apply and labour intensive (Xu et al. 2018). The relatively high cost of obtaining these data can be justified only for major transmission water mains where the cost of failure is high (Alvisi and Franchini 2010; Nishiyama and Filion 2013). Therefore, the physical predictive models have been excluded from the literature review since they are beyond the scope of the study.

Individual pipes vs aggregated models

Kleiner and Rajani (2012) concluded that due to the inherent uncertainty and the lack of enough data is not feasible to reliably analyse the failure pattern of a single pipe. Therefore, most predictive models (e.g. Shamir and Howard 1979; Lei and Sægrov et al. 1998; Le Gat and Eisenbeis 2000; Pelletier et al. 2003; Berardi et al. 2008; Asnaashari et al. 2009) have been developed on a group level to reduce the complexity of the rehabilitation problem (Roshani and Filion 2013) and conduct more effective analysis assuming that pipes with share the same characteristics are expected to have a similar failure rate (Kleiner and Rajani 2012). The *aggregation* criteria that have been used so far include pipe properties (e.g. material, diameter and age), external factors (e.g. soil type), geographical clustering and the number of previous failures (Osman and Bainbridge 2010). Two conflicting goals should be simultaneously fulfilled in the pipe *aggregation* process. The homogenous groups should be small enough to be uniform and large enough to obtain models with a meaningful ‘goodness of fit’ (Kleiner and Rajani 2001).

2.3.1 Statistical Models

The statistical models link pipe breakage patterns to various pipe descriptive variables and other environmental and operational factors using regression analysis of historical pipes break data (Kleiner and Rajani 2001) and are applicable to various levels of input data. They can cope with the lack of sufficient knowledge related to the complex physical mechanisms that lead to pipe failure and are divided into deterministic and probabilistic models (Kleiner and Rajani 2001). Probabilistic models can cope with randomness and probabilities, while deterministic models deal with 'crisp' data without any presumed randomness (Kleiner and Rajani 2001). The outcome of a probabilistic model is a single probability or a set of probabilities (e.g. probabilities distribution) whereas the outcome of a deterministic model is a certain value (e.g. number of failures or failure rate) (Dehghan et al. 2009).

Kleiner and Rajani (2001) described exhaustively the statistical approaches developed prior to their review. Hence, these approaches have been excluded from the literature review. Nishiyama and Filion (2013) reviewed the statistical models and partly the data-driven models developed between 2002 and 2012.

Deterministic Models

Statistical deterministic models (Table 2.2) where the first developed predictive models and are further divided into exponential and linear regression (Kleiner and Rajani 2001). The regression parameters of these models are considered fixed and least-squares estimation or maximum likelihood methods are utilized to determine the regression parameters or coefficients (Kleiner and Rajani 2002; Boxall et al. 2007; Asnaashari et al. 2009).

Table 2.2 Summary of Deterministic Models

References	Variables	Model
Kleiner and Rajani (2000)	Ageing, Freezing Index, Rain Deficit Cumulative Length of Replaced Mains, Cumulative Length of Cathodic Protection	$MERM^a$
Kleiner and Rajani (2002)	Freezing index, rainfall deficit, snapshot Rain deficit, length of replaced mains, cumulative length of cathodic protection	$MERM^a$
Boxall et al. (2007)	Length, diameter, age, material, soil corrosivity	$PGLRM^b$
Asnaashari et al. (2009)	Length, diameter, wall thickness, maximum pressure, pipe location, cover depth, previous breaks	$MLRM^c$ $PGLRM^b$
Wang et al. (2009a)	Length, age, size	$MLRM^c$
Yamijala et al. (2009)	Length, diameter, age, pressure, time Since last break, temperature, rainfall, land use, soil type, min-max soil moisture, soil corrosivity	$MLRM^c$ $MERM^a$ $PGLRM^b$ $LGLM^d$

[a]MERM=Multivariate Exponential Regression Model

[b]PGLRM=Poisson Generalized Linear Regression Model

[c]MLRM=Multivariate Linear Regression Model

[d]LGLM=Logistic Generalized Linear Regression Model

Exponential regression models

The most general form of the exponential regression models (ERM) is given as:

$$Y=\mathrm{f}(\vec{x};\vec{\beta}) \quad (2.1)$$

Where Y is the dependent variable, the function $\mathrm{f}(\vec{x};\vec{\beta})$ is non-linear with respect to the unknown parameters $\beta_0,\ \beta_1,\ldots.\ \beta_n$ and $x_0,\ x_1,\ldots.\ x_n$ are the explanatory variables

Shamir and Howard (1979) were the first to suggest that water mains breakage rates increased exponentially with pipe age. This single variate, two-parameter expression has been used by others with and without modifications, (e.g. Walski and Pelliccia 1982; Clark et al. 1982; Kleiner et al. 1998; Kleiner and Rajani 1999).

Kleiner and Rajani (2000) proposed a generalization to a multi-variate exponential model (MERM) (Eq. 2.2) to analyse breakage rate patterns of water mains, considering ageing, environmental (Freezing Index and Rain Deficit) and operational (Cumulative Length of Replaced Mains and Cumulative Length of Cathodic Protection) factors as explanatory variables.

$$N_{(x)}=N_{(x_0)}e^{\underline{a}\underline{x}T} \quad (2.2)$$

Where $N_{(\chi)}$ is the number of breaks, $\underline{x}$ is the vector of time-dependent covariates, a is the vector of parameters corresponding to the covariates, x_0 is the value of the covariates at a baseline time

The method was demonstrated with three case studies: for CI and AC pipes in Adelaide (Australia), CI pipes up to 12'' diameter in Ottawa (Canada) and CI in Edmonton (Canada). Adelaide's climate is warm and arid, thus only the impact of time and Rain Deficit was examined. All the covariates were used in Ottawa's case, whereas Cathodic Protection was excluded in Edmonton's case. The coefficient of Determination (R^2) was 0.44 for the CI pipes and 0.70 for the AC pipes in Adelaide's case. The obtained accuracy was slightly higher (R^2=0.78) for the Ottawa's case while for Edmonton's case study was significantly improved (R^2=0.86). The accuracy of the predictions for both CI and DI pipes in the Adelaide's case study is not satisfactory.

Kleiner and Rajani (2002) used the MERM proposed by Kleiner and Rajani (2000) for the short-term (3 to 4 years) and long-term prediction of water main breaks considering aging, environmental and operational factors. The proposed method was demonstrated to the CI and DI pipes with up to 12'' diameter pipes in six case studies in Ontario (Canada). The first case study considered only aging, while CP, cumulative RD, FI, snapshot of rainfall deficit and the length of replaced mains were included gradually in the analysis of the remaining five regions (i.e. in the first case study only one explanatory variable was considered whereas all of them were considered in the sixth case study). The accuracy for the six case studies was evaluated using the R^2 and the adjusted coefficient of determination (R_a^2). The values of R^2 range from 0.619 (case 1) up to 0.793 (case 6) constantly increasing. The inclusion of more explanatory variables increased the accuracy of the predictions. The values of R_a^2 increase from 0.603 (case 1) up to 0.736 (case 4) and remained constant in case 5 with a small decline in case 6 (0.731). The decline of R_a^2 in case 6 can be attributed to the fact that the water utility implemented an aggressive main replacement program during the examined period. The inclusion of a period in which breakage rates predominantly decrease, cause the MERM to yield results that are counterintuitive, such as the negative effects of mains replacement program (e.g. Kleiner and Rajani 2002). The authors expressed their concerns for the parallel investigation of the failure patterns of two different materials (i.e. CI and DI). The MERM used by Kleiner and Rajani (2002) did not provide any further improvement in the accuracy of the predictions. Both approaches are mathematically simple to understand, but with moderate accuracy because they involve a great deal of conditional assumptions, specified data and statistical analysis (Kleiner and Rajani 2001) and therefore there is a limited number of implementations the last few years.

Linear regression models

Linear regression models (LGM) in their simplest form assume that the variable of interest Y is a linear function of a set of explanatory variables X_i given as:

$$Y=b_0+\sum_i b_i X_i \quad (2.3)$$

Where Y is the dependent variable, b_0, b_i are the unknown constants to be estimated and X_i is the set of explanatory variables

The use of a linear relationship was first suggested (Kettler and Goulter 1985). Generalized linear models (GLMs) have extended the linear regression to allow for discrete data. Thera are two types of GLM, Poisson GLM and logistic regression GLM.

A regression model based on the Poisson distribution counts on the observed values of the covariates and specifies that the conditional mean of the counts is given by a continuous function $\mu(\beta,\vec{x_i})$ of the covariates values as given by Eq. 2.4 where β is the n x 1 vector of the regression parameters:

$$E[y_i|\vec{x_i}]=\mu(\vec{\beta},\vec{x_i}) \quad (2.4)$$

Where $\mu(\vec{\beta},\vec{x_i})$ is a continuous function, x_i is the vector of n covariates for system segment i and the number of failures on segment i be given by y_i and β the n x 1 vector of regression parameters

Logistic regression GLM predicts the probability of a discrete outcome from a set of explanatory variables that may be discrete, continuous, or dichotomous or a combination of these (Yamijala et al. 2009). The dependent variable in a binary logistic regression model takes the value 1 with a probability of P or the value 0

with a probability of (1-P). the probability of the binary event is modelled using a logit transformation of P as:

$$P=\frac{\exp[\alpha+\beta_1\chi_1+\beta_2\chi_2+\cdots+\beta_i\chi_i]}{1+\exp[\alpha+\beta_1\chi_1+\beta_2\chi_2+\cdots+\beta_i\chi_i]} \quad (2.5)$$

Where a is the constant regression parameter, β_i are the regression coefficients for the explanatory variables and X_i are the independent variables

With this model, the log of the odds ratio is linear in the explanatory variables:

$$\text{Log it}[P(x)]=\log[P(x)/(1-P(x))]= \alpha + \beta_1\chi_1 + \beta_2\chi_2 + \cdots + \beta_i\chi_i \quad (2.6)$$

Wang et al. (2009a) used MLRM to forecast the annual break rate (R) of individual water mains considering pipe age (A), length (L), depth of installation (S) and any combination of them as explanatory variables. An individual model was developed for each material: grey CI (Eq 2.7), DI without lining (Eq. 2.8), DI with lining (Eq. 2.9), PVC (Eq. 2.10) and concrete (Eq. 2.11), using Minitab statistical software.

$$\text{Log}_{10}R=4.85-0.0206A+0.000245A^2+0.00281S-0.905\text{Log}_{10}L-1.40\text{Log}_{10}S \quad (2.7)$$

$$\text{Log}_{10}R=1.83-0.911\text{Log}_{10}L \quad (2.8)$$

$$\text{Log}_{10}R=3.36+0.000150(L*A)-1.11\text{Log}_{10}L-0.646\text{Log}_{10}A-0.254\text{Log}_{10}S \quad (2.9)$$

$$\text{Log}_{10}R=2.69-0.898\text{Log}_{10}L-0.745\text{Log}_{10}A \quad (2.10)$$

$$\text{Log}_{10}R=1.81+0.00593L-0.000028(L*S)-0.958\text{Log}_{10}L \quad (2.11)$$

The $R^2(\%)$ of the models is 68.9, 65.0, 71.5, 78.9 and 81.3 for the grey CI, DI without lining, DI with lining, PVC and concrete pipes respectively. Their models have minimal data requirements but as indicated by the low values of R^2, they need improvement (Wang et al. 2009b). This approach is the only MLRM applied on an individual pipe resulting in very low accuracy possibly due to data scarcity.

Asnaashari et al. (2009) compared the ability of a MLRM (Eq. 2.12) and a PGLRM (Eq. 2.13) to predict the failure frequency (FF) of water mains. The pipes were divided into four groups based on the material (AC, CI, DI and PE). The candidate variables were pipe diameter (DR), pipe wall thickness (TK), cover depth (DP), pipe age (PA), pipe length (LL), maximum pressure (MP), pipe location (PL) and failure history (FH) and the dependent variable was the Failure Frequency (FF) during a 10 years period.

$$FF=a_0+a_1DR+a_2LL+a_3DP+a_4TK+a_5MP+a_6AG+a_7PL+a_8FH \quad (2.12)$$

$$FF=\exp(b_0+b_1DR+b_2LL+b_3DP+b_4TK+b_5MP+b_6AG+b_7PL+b_8FH) \quad (2.13)$$

Where the regression coefficients a_0, a_1, …a_8 and b_0, b_1, …b_8 were determined by the degree of their contribution to FF

The R^2 of the obtained multiple regression models was 0.77, 0.52, 0.88 and 0.69 for the AC, CI, DI and PE respectively. The R^2 of the multiple Poisson models was 0.79, 0.71, 0.95 and 0.75 respectively. The Poisson model had superior prediction capabilities (higher values of R^2) because it can handle non-linear relationships and independent pipe variables, thus addressing the fitting problems, but suffered from over-dispersion problems. Despite the improvement, the accuracy is moderate for most materials (expect DI pipes).

Boxall et al. (2007) also used a Poisson generalised linear model (Eq. 2.14) to analyse the relationship between burst rate and age, diameter and length for CI and AC pipes in the UK.

$$\gamma(D,L,A)=\alpha+\beta_D D+\beta_L L+\beta_A A \quad (2.14)$$

Where D is the diameter, L is the length, A the age and a, β_D, β_L, β_A the coefficients to be estimated

A slightly curved relationship was found between burst rate and length (explained by considering length as a surrogate for connection density and joints which can be points of potential weakness). The relationship between burst rate and diameter was found to increase exponentially with decreasing diameters while the relationship between burst rate and age was complex (direct for the AC and indirect for the CI), probably due to age acting as a surrogate for other explanatory variables. The variation in annual burst rate due to different soil properties is small compared to the other explanatory variables, suggesting a relative lack of dependence between soil and burst rate. Boxall et al. (2007) showed the prediction abilities of a PGLRM on a group level analysis.

Yamijala (et al. 2009) included all types of deterministic models in their analysis. They developed a time linear model (Eq. 2.15), a time exponential model (Eq. 2.16), a Poisson GLM ((Eq. 2.17) and a logistic GLM model (Eq. 2.18) to estimate the number of breaks for a six-month period. The explanatory variables were: diameter (D), pipe segment length (L), pipe material (AC, CI, concrete steel cage, DI, PVC, steel), the year of installation (INSTYR), the time since the last break (TIME), the operating pressure of the pipe (PRE), the land use above the pipe ($LU_1 - LU_{11}$), the type of the soil around the pipe ($ST_1 - ST_4$), the temperature (TEMP), the rainfall (RAIN), the maximum soil moisture (SMAX), the difference between maximum and minimum soil moisture (MX-MN) and three principal components ($PC_1 - PC_3$) related to six soil corrosivity covariates. The time linear, the time exponential and the Poisson model were fit in R and were applied only to pipes that had experienced at least a break during the data-recording period to avoid a zero-inflation problem. The logistic GLM was applied to all pipe segments.

Y=-0.0027D-0.44AC-0.45CI-0.34CSC-0.46DI-0.45PVC+2.6x10^{-5}L-0.00027LU_6-0.000327LU_8-0.00035LU_{11}+0.0018TEMP+3.7x10^{-5}RAIN+0.0015SMAX (2.15)

Y=-0.3EXP[0.47TIME+2.7x10^{-5}INSTYR] (2.16)

log_μ=-6.84-(0.023)(D)+(0.12)(AC)+(2.2x10^{-4})(L)-(2.6x10^{-3})(LU_6)-(3.2x$10^{-3)}$(LU_8)-(3.3x10^{-3})(LU_{11})+(0.0166)(TEMP)+(2.65x10^{-4})(RAIN) (2.17)

Logit[P(x)]=log[P(x)/(1-P(x))]=-5.82-0.12(D)+(0.02)(LU_1)+(0.02)(LU_2)+(0.02)(LU_3)+(0.02)(LU_4)+(0.02)(LU_5)+(0.01)(LU_6)+(0.02)(LU_7)+(0.02)(LU_9)+(0.02)(LU_{11})-(0.01)(ST_1)-0.005(ST_3)-(0.02)(PC_1) (2.18)

The tests show that the time linear, time exponential and the Poisson GLM models do not fit particularly well while the logistic was fitted well with the data because it accounts for the excess zero presence in a way that the other models do not. The time linear, time exponential and the Poisson GLM models were applied only to pipes that had experienced at least a break during the data-recording period limiting their applicability because in most networks only a small fraction of pipes has at least a recorded failure. Also, their ability to cope with the zero-inflation problem can't be evaluated. Yamijala et al. (2009) showed that the LGLM outperforms the other types of probabilistic models because it can effectively cope with the zero-inflation problem.

Probabilistic models

Probabilistic models (Table 2.3) are further divided into Bayesian, Proportional hazards, Accelerated life-time, Poisson and Yule Process (Kleiner and Rajani 2001; Park et al. 2011; Economou et al. 2012; Rajani et al. 2012; Martins et al. 2013; Kabir et al. 2015b).

Table 2.3 Summary of Probabilistic Models

References	Variables	Model
Watson et al. (2004)	Age	BBN^a
Park et al. (2008)	Break history, material, installation time	$ROCOF^b$
	diameter, length	PLP^c
Park et al. (2011)	Land development, internal pressure,	PHM^d
	length, number of customers	
Economou et al. (2012)	Length	$NHPP^e/ZINHPP^f$
Rajani et al. (2012)	Weather and air temperature-related	$NHPP^e$
Martins et al. (2013)	material, diameter, length, installation	$PM^g/WALM^h$
	year, number of previous failures	$LEYP^i$
Kabir et al. (2015b)	Age, length, diameter, soil resistivity	BBN^a
	soil corrosivity, vintage	
Kabir et al. (2015c)	Failure rate, age, diameter, length,	BBN^a
	rain deficit, average temperature, freezing	
	index, land use soil resistivity, vintage	
Kabir et al. (2015d)	Age, length, diameter, vintage, soil	$BMAM^k$
	Resistivity soil corrosivity, temperature,	$BWPHM^l$
	freezing index, rain deficit, connections	
Kimutai et al. (2015)	Length, diameter, material, soil resistivity,	$WPHM^m$
	Freezing index, rain deficit	PHM^d/PM^g

[a]BBN=Bayesian Belief Network

[b]ROCOF=Rate of Occurrence of Failure

[c]PLP=Power Law Process

[d]PHM=Proportional Hazard Model

[e]NHPP=Non-Homogenous Poisson Process

[f]ZINHPP=Zero-Inflated Non-Homogenous Poisson Process

[g]PM=Poisson process Model

[h]WALM=Weibull Accelerated Lifetime Model

[i]LEYP=Linear Extension Yule Process

[k]BMAM=Bayesian Model Averaging Method

[l]BWPPM=Bayesian Weibull Proportional Hazard Model

[m]WPHM=Weibull Proportional Hazard Model

Proportional hazards models

The Proportional hazards models (PHMs) model the time-dependent ageing process of a pipe assuming a multiplicative relationship between them (Park et al. 2008). They are divided into Cox's PHMs, Rate of Occurrence of Failure (ROCOF) and Power Law process (PLP) models (Kleiner and Rajani 2001).

The PHM was initially developed by Cox (1972) and its general form is:

$$h(t,Z)=h_o(t)e^{bTz} \quad (2.19)$$

Where h(t,Z) is the hazard function which is the instantaneous rate of failure (probability of failure at time t+Δt given survival to time t), $h_o(t)$ is the arbitrary baseline function, Z is the vector of covariates acting multiplicatively on the hazard function, b is the vector of coefficients to be estimated by regression from available data

The log-linear ROCOF and the PLP are expressed respectively, as:

$$v(t)= \exp(\beta_t + \beta_1 t) \quad (2.20)$$

$$v(t)= \gamma\delta t^{\delta-1} \quad (2.21)$$

Where $v(t)$is the failure intensity and $\beta_t, \beta_1, \gamma, \delta$ the parameters to be estimated

Accelerated lifetime models

Accelerated lifetime models are multivariate models in which the logarithm of the time to next failure is defined as the linear combination of covariates x=[$x_1 x_2 x_p$] and a random error variable Z:

$$\ln T=x^T\beta + \sigma Z \quad (2.22)$$

Where β=[$\beta_0 \beta_1 \ldots\ldots \beta_p$] are the unknown parameters and σ is a scale parameter

Single-variate Poisson Process

A Poisson process counts the number of failures {N(t),t≥0} within a time interval of zero and t (0,t] with rate γ, where N(t) ∈ N_0=number of occurrences during t, satisfying the following conditions (Røstum 2000):

1. N(t)>0
2. N(t) is an integer
3. If $t_1 < t_2$ then, N(t_1)≤N(t_2); and
4. For t_1<t_2, [N(t_2)-N(t_1) represent the number of failures that have occurred in the time interval [t_1, t_2]

It follows that if the number of random events N(t) is Poisson distributed, then the probability of the occurrence of n failures P{N(t)=n} ng is estimated as:

$$P\{N(t)=n\}=\frac{e^{-\pi t}(\lambda t)^n}{n!} \quad (2.23)$$

Where n=1, 2 represents the number of observed pipe breaks and λ is the coefficients of covariates, which represents the rate of the Poisson process

Linear Yule Process

The Linear Yule Process (LEYP) is a counting process in which the intensity function is a linear function of the number of past events and depends on the age of the process. Le Gat (2009) demonstrated its probability function as:

$$P\{N(t)-N(s)=n|N(b)-N(a)=j\}=\frac{\Gamma(a^{-1}+j+n)}{\Gamma(a^{-1}+j)n!}\frac{[\mu(b)-\mu(\alpha)+1]^{a^{-1}+j}\,[\mu(t)-\mu(s)]^n}{[\mu(t)-\mu(s)+\mu(b)-\mu(a)+1]^{a^{-1}+j+n}} \quad (2.24)$$

Where a is the parameter associated with the number of previous events, λ(t) is the function that translates the aging and the covariates effect, $\mu(t)=e^{\alpha\Lambda(t)}$ and $\Lambda(t)=\int_0^t \lambda(u)du$

Eq. (2.25) implies that the distribution of failures of LEYP is a continuous extended Negative Binomial:

$$\{N(t)-N(s)|N(b)-N(a)=j\}\text{-NB}\left(a^{-1}+j,\frac{\mu(b)-\mu(a)+1}{\mu(t)-\mu(s)+\mu(b)-\mu(a)+1}\right) \quad (2.25)$$

One important feature of a failure prediction model is that it should distinguish the probability of failure in different pipes by their different attributes. Thus, the intensity function based on pipe covariates is transformed as Le Gat (2009):

$$\lambda(t)=\delta t^{\delta-1}e^{x^T\beta} \quad (2.26)$$

Where δ is the aging factor of pipes and β is the vector of parameters associated with the covariates x

Bayesian Belief models

Bayesian Belief Network (BBN) is a graphical model that permits a probabilistic relationship among a set of variables (Pearl 2014). BNN is based on the Bayes' theorem that manages uncertainty by explicitly representing the conditional probability dependencies between variables (Tang and McCabe 2007; Cinar and Kayakutlu 2010). In a BNN analysis, for n number of mutually exclusive parameters $X_{i(i=1,2,\ldots,n)}$ and a given observed data Y, the updated probability is computed as (Peral 2014):

$$p(X_i|Y)=\frac{P(Y|X_i)\ \text{x}\ P(X_i)}{\sum_j p(Y|j)p(X_{j)}} \quad (2.27)$$

Where p(X|Y) represents the posterior occurrence probability of X, given the condition that Y occurs, p(X) denotes the prior occurrence probability of X, p(Y) denotes the marginal occurrence probability of Y and is effectively constant since the obtained data is in hand and p(X/Y) refers to the conditional occurrence probability of Y, given that X occurs too

Park et al. (2008) applied the log-linear ROCOF and the Power Law process to estimate the optimal replacement time considering, break history, material, installation time, diameter and length as explanatory variables. The estimated replacement time is optimised and modelled by applying the methodology outlined by Loganathan et al. (2002) as:

$$Brk_{th}=\frac{ln[(1+R)/(1+i)]}{ln[1+(\frac{C}{F})]} \quad (2.28)$$

Where R is the annual interest rate (1/year), *i* is the annual inflation rate (1/year), C is the repair cost of a break and F is the replacement cost of a pipe

The comparison of standard deviation showed that the log-linear ROCOF was a better model over the power law process when the failure time model was used. This conclusion indicates that recording each failure time and applying this to models is more accurate than using failure numbers over a time interval. Both methods require a sufficiently large number of recorded breaks due to the fact that are intended for the non-linear modelling of the failure rates. Therefore, some limits of recorded number of breaks should exist and in line with this requirement, at least five numbers of failures are assumed for this study. The model needs a sufficiently large number of break records which is not always available and can't assess the effects of various factors that cause pipe failure.

Park et al. (2011) attempted to address this problem by adding extra information such as failure-related effects, general conditions and survival probabilities. They constructed a PHM for the time intervals between consecutive pipe breaks using the degree of land development (DL), internal pressure type of pipe (PT), length of pipe (L) and the number of customers in a grid (C) as explanatory variables. In addition, material-joint types were also considered as covariates for pit-CI, spun

CI rigid joint (SR), and spun CI flexible joint (SF) types. Individual pipes were allocated into seven groups according to the past break history (from 0 up to 6 previous breaks). The main shortcoming of this type of models is that they require an extended database collected in a standardised framework which reduces their applicability since most of the water utilities record pipe failure in their distinctive way. Park et al. (2011) stated that the implementation of the PHM may require other types of data and some coding with a high-level computer language that may not be always available.

Economou et al. (2012) compared the ability of a Bayesian Non-Homogenous Poisson Process (NHPP) model and a Bayesian zero-inflated Non-Homogenous Poisson Process (ZINHPP) model to handle the excess amount of zeros in the number of failures (zero-inflation problem). The proposed methodology was applied to a set CI pipes in North America and a set of AC pipes in Australia. The available explanatory variable for the CI pipes was only the length. The available explanatory variables for the AC pipes were pipe length, pipe diameter, maximum absolute pressure and maximum pressure change within 24hrs. In both cases, the ZINHPP model fitted the data better than the NHPP model for the calibration dataset. It can be attributed to the fact that the ZINHPP accounts for the possibility of more zeros in failure counts than would be expected from the NHPP alone, by allocating extra probability to the possibility of no failure (Santos et al. 2017). The main limitations of the approach are that length is the only available variable for the CI pipes and many other important variables are missing while the observation period for the AC pipes was too short and therefore, no validation period was considered for assessing the generalization capabilities of the approach.

Rajani et al. (2012) also used a non-homogenous Poisson-based (NHPP) pipe deterioration model to examine the impact of air temperature-based and water temperature-based covariates on breaks of homogenous groups of CI, DI and galvanized steel pipes for time steps with varying duration (lasting between 5 and 90 days). Short time steps capture better the temperature fluctuation, but the large number of data, introduces a lot of 'noise' making difficult to obtain mathematical relationships with meaningful 'goodness of fit'. On the contrary, long time steps require careful selection of a starting point in order not to miss seasonal temperature variations and they result in fewer data which may lead to over-fitting. The trials were evaluated with the R^2 and the likelihood ratio test. These performance indicators were used to identify the covariates with the highest impact. A modified form of the Poisson model proposed by Kleiner and Rajani (2010) was used here, where the probability of $P(k_i)$ observing k_i breaks in time step i, in terms of one or more time-dependent covariates, is:

$$P(k_i)=\frac{\lambda_\iota^{\kappa_\iota}\exp(-\lambda_i)}{\kappa_\iota!} \quad (2.29)$$

$$\lambda_i=\exp\left[\beta_o+\psi\tau(g_i)+\beta q_i\right] \quad (2.30)$$

Where λ_i is the expected number of breaks (or the rate of occurrence of breaks) in time step i, β_o is a constant, q is a row vector of time-dependent covariates prevailing at time step i and β is a column vector of the corresponding coefficients to covariates q

Average mean air temperature, maximum air temperature increases and decreases and how fast the air temperature increases and decreases (intensities) over a specific period of days, were identified as the most significant covariates. Based on the aforementioned arguments and the evaluation of the accuracy the time step of 30 days was chosen. The proposed methodology showed a good

accuracy but lacks in validation on a test dataset since the main objective was to identify the influence of temperature-based covariates rather to be used for forecasts.

Martins et al. (2013) compared the ability of a single-variate Poisson process (PM), a Weibull accelerated lifetime model (WALM), and a linear-extended Yule process (LEYP) model to identify the pipes (AC, high-density PE and PVC) that are most likely to fail using short failure records. The available pipe variables were material, diameter, length, installation year, and number of previous failures. The single-variate Poisson process used as grouping criteria the pipe material, three groups of pipe diameter, and three groups of pipe length. The installation year was disregarded because of the high correlation between this variable and pipe material. The age and the number of previous failures were disregarded from the explanatory variables. Pipe material was used as a grouping criterion in the WALM. The logarithm of length, the diameter, the previous failures and the pipe age at the last recorded failure entered the regression model as covariates in the WALM. The linear extended Yule process used pipe material as a grouping criterion and considered the logarithm of length and pipe diameter as covariates. The pipe age and the number of previous failures were already considered during model construction, and thus were not used as covariates. Two methods of defining the training and test sets were considered: the temporal division and the random division methods. For the temporal method, the most recent three years of data are reserved for model validation while for the random division method, 50% of pipes are selected at random to be used for training, and the remaining 50% is used for validation. The WALM yielded the best results among the three models, however, it is based on Monte Carlo simulations, which can be time-

consuming. The distribution of the number of failures during a period using the WALM can't, however, be analytically derived, because the convolution of Weibull distributions cannot be analytically calculated. In WALM and LEYP it is assumed that the Weibull scale parameter and the process rate, respectively, are proportional to the exponent of a linear combination of the covariates vector. The LEYP effectively detected pipes with higher failure likelihood, but tends to overestimate the number of future failures, probably due to the linear increment of the intensity function with the number of previous failures. The single-variate PM is the simplest but the least accurate.

Kabir et al. (2015b) developed a Bayesian multiple regression based on data fusion model to predict the failure of CI and DI pipes in Calagary's (Canada) WDN. The explanatory variables of the model for the CI pipes were age, length, diameter, vintage, soil resistivity, while the explanatory variables for the DI model were age, length, diameter, vintage and soil corrosivity index. Separate models were developed for small diameter pipes (≤150mm) and big diameter pipes (>150mm). This method is designed to merely analyse individual causes, instead of combined causes. Moreover, it can be difficult for domain experts to elicit the casual Bayesian structure with combined causes from domain knowledge only (Ma et al. 2016).

Kabir et al. (2015c) compared the accuracy of a MLRM and the Bayesian multiple regression model developed by Kabir et al. (2015b) in estimating the number of breaks. The proposed methodology was implemented to the CI and DI pipes that had experienced at least one break during the data-recording period. Both models were developed using the software R, considering number of previous

factors, break/failure rate, pipe age, pipe diameter, pipe length, average temperature, freezing index, rain deficit, vintage, land use and soil resistivity as explanatory variables. The physical factors were found to be the most significant followed by the environmental factors. Between the environmental factors, FI and rain deficit had a higher impact on CI rather than DI pipes. The Bayesian multiple regression model is more accurate than the MLRM because it can consider multiple information from different sources and combine them considering their interdependencies resulting from cause-effect characteristics. The main drawback is that the case study entails only pipes that have experienced at least one break reducing significantly the size of the sample.

Kabir et al. (2015d) used a Bayesian model averaging method (BMA) and a Bayesian Weibull Proportional Hazard Model (BWPHM) to develop survival curves and predict the failure rates of CI and DI pipes. The explanatory variables were pipe age, pipe length, pipe vintage, pipe diameter, soil resistivity, soil corrosivity index, temperature, FI, rain deficit and the number of connections of each pipe. Both CI and DI pipes were split into groups based on whether the Number of Previous Failures (NOPF) was zero or above zero. The obtained results showed a different response for the CI and DI pipes to the effect of covariates. The results also represented that the survival times of CI and DI pipes with NOPF=0 are higher than NOPF>0. After the first break, soil resistivity is the most significant or influential parameters for the increases the hazard of the CI pipes whereas DI pipes are more sensitive on soil corrosivity index.

Kimutai et al. (2015) compared a Weibull proportional hazard model (WPHM), a Cox proportional hazard model (Cox-PHM), and a Poisson model (PM) in

describing the effects of physical and environmental factors on the failure of CI, DI and PVC pipes. Results from the preliminary process indicated that covariates with the most significant impact in influencing pipe breaks are length, diameter, pipe material, soil resistivity, soil resistivity, FI and RD. The physical covariates were found to be more critical while environmental covariates had an impact only on PVC pipes. The WPHM captured all the breaking phases of the metallic pipes better than the Cox-PHM and the PM which underestimated the number of breaks. The accuracy of Cox-PHM decreased as the pipe shifted from slow to fast breaking phases of pipe life while PM's increased. For PVC pipes the performance of PHMs in the prediction of these pipes was very low, which could be attributable to time dependent hazards and low number of breaks observed. The Cox-PHM assumes a proportional fixed effect on the baseline hazard function which depends on the time but not on the covariates (Alvisi and Franchini 2010) and represents the aging process such as the effect of internal and external corrosion (Clark et al. 2010), which occurs not only as a function of time but also other stressing variables (Le Gat and Eisenbeis 2000). The Cox-PHM is difficult to get good breakage risk estimates without a large dataset, which is the case for many water utilities.

Overall, the statistical models that predict the behaviour of water pipes are affected by both the quantity and quality of available data (Díaz et al. 2016; Gómez-Martínez et al. 2017), and by the applied statistical techniques (e.g. selection of probability distribution function) (Boxall et al. 2007; Economou et al. 2007, 2008). Soft computing is viewed as an alternative method to hard or precise computing, in that it is tolerant of uncertainty and imprecision (Nishiyama and Filion 2013).

2.3.2 Data-Driven Methods

The data-driven methods can handle imprecision, missing or partial data and simulate complex non-linear relationships between inputs and outputs for large systems (Fayyad et al. 1996; Giustolisi and Savic 2006; El-Baroudy et al. 2010). They can be classified based on the level of prior information required in the construction phase as white-box, grey-box and black-box models (Giustolisi 2004; Giustolisi and Savic 2006; Giustolisi et al. 2007; Fiore et al. 2012; Nishiyama and Filion 2014).

- A white-box model (e.g. physical models) is a system where all necessary information is available and is based on physical laws and known variables and parameters. Because the variables and parameters have physical meaning, they also explain the underlying relationships of the system. However, model construction can be difficult if the underlying mechanisms are not fully understood or the experimental results obtained in the laboratory environment do not correspond well to the prototype environment.
- A black-box model (e.g. ANN) is a system in which there is no prior information is available. The functional form of relationships between variables and the numerical parameters in those functions are unknown and need to be estimated.
- Between the white-box and the black-box models there is a wide range of grey-box models (e.g. EPR). Their mathematical structure is derived through the conceptualisation of physical phenomena or through the simplification of differential equations, describing the phenomena under consideration. Grey-box models usually estimate parameters by means of input/output data analysis, although some information about the underlying relationship is normally already known.

Artificial Neural Network models

Artificial Neural Network (Table 2.4) are data-driven 'black-box' models that can capture the complex relationship between input and output pipe failures by learning from historical data even in the absence of physical consideration (Giustolisi and Laucelli 2005; Al-Barqawi and Zayed. 2008; El-Baroudy et al. 2010; Rezania et al. 2010). They have been a good modelling approach for pipe break forecasting and particularly well suited to handle large data sets and multiple variables (Park et al. 2008). Typically, an ANN model consists of an input layer which receives the inputs. Through hidden layers these inputs are processed to provide the output layer with the values predicted by the network.

Table 2.4 Summary of Artificial Neural Network models

References	**Variables**
Ahn et al. (2005)	max and min soil temperature 1.5m underground, min and max water temperature, max and min air temperature
Achim et al. (2007)	Length, diameter, age, material, geographical coordinates
Tabesh et al. (2009)	Length, diameter, age, depth of installation, average pressure
Jafar et al. (2010)	Length, diameter, age, material, soli type, location in the street, pressure, protection
Asnaashari et al. (2013)	Length, diameter, age, material, break category, year of cement-mortar lining and cathodic protection (if done), soil type

Ahn et al. (2005) developed an ANN model to predict pipe breaks considering environmental factors (i.e. soil temperature at 1.5m under ground; maximum and minimum values of water temperature; maximum and minimum values of atmosphere temperature), the ratio of stainless steel, galvanized steel CI and

ductile CI pipes and the ratio of metered water. Most pipes (except CI) did not have an increased number of failures during the colder winter months, but in spring and fall when the water temperature fluctuates and soil shrinks/swells. The rapid increase and decrease of temperature were found to be the most significant factors. The ANN model had a good performance in predicting the pipe breaks on a seasonal basis. However, the sensitivity analysis indicated a low accuracy in case of rapid increase and decrease of pipe failure frequency. This ANN model is the only one that considers environmental factors as input.

Achim et al. (2007) used a multi-layer ANN to predict the number of failures/kilometres/years for individual CI and CI concrete lined pipes. The performance of the ANN model was compared with a shifted time power model and a shifted time exponential model. Physical factors (i.e. diameter, year of construction, age, length) and the pair of geographical coordinates were considered as explanatory variables. The ANN model outperformed statistical models where databases were relatively large and noisy. However, the values given for fit were not very satisfactory (Moselhi and Fahmy 2007) which can be attributed to the fact that the analysis was conducted on an individual pipe level.

Tabesh et al. (2009) developed an ANN and a neuro-fuzzy model to predict pipe failure rate using pipe diameter, length, age, depth of installation and average hydraulic pressure as explanatory variables. Then a multivariate regression method was used to evaluate the performance of the two data-driven models. The proposed methodology was demonstrated by implementation in the steel, CI and AC pipes of a WDN in Iran. The ANN model provided the most realistic and accurate results in predicting the pipe failure rate.

Jafar et al. (2010) developed six distinctive ANN models to predict the failure rate of AC, PE, and metallic pipes, all pipes, pipes with low number of failures and pipes with high number of failures respectively. The included physical (material, length, diameter, thickness, and age), environmental (type of soil, location in the street) and operational (pressure and cathodic protection) factors. Very high accuracy $(R^2 = 0.972)$ was obtained for the model with the ''high number of failures'' pipes. The usefulness of this model is relatively limited since very few pipes in a WDN have a recorded high number of failures. The rest models showed a significantly lower accuracy.

Asnaashari et al. (2013) compared the accuracy of an ANN model and a MLRM in forecasting the failure of CI and DI pipes with diameter 25, 37, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500 and 600 mm. The explanatory variables considered were: pipe length, diameter, age, break category, soil type, pipe material, the year of Cement Mortar Lining and CP (if implemented). The ANN model $(R^2 = 0.94)$ outperformed the MLRM $(R^2 = 0.75)$ indicating that is more successful in predicting failure rates since it can simulate the non-linear relationship between pipe failure the factors that cause it.

The shortcomings of the ANN are that they require the structure of a neural network to be identified and the initial set-up can be time-consuming and complicated (Giustolisi and Laucelli 2005) and they are prone to overfitting (Giustolisi and Laucelli 2005; Fahmy and Moselhi 2009). There is concern that the ANN may be 'over-trained' resulting in a model that is just capable of 'memorising' the training data set rather than being able to generalize the patterns to new sets.

Genetic Programming Models

Genetic programming (GP) is an evolutionary algorithm-based method that mimics the natural evolutionary selection and allows a global exploration of the space of model expressions. The symbolic regression manipulates populations of solutions using operations analogous to the evolutionary processes that operate in nature. The genetic programming procedure mimics natural selection as the 'fitness' of the solutions in the population improves through successive generations. The symbolic regression GP generates 'transparent' models allowing to gain understanding of the relationship between failure and the explanatory variables. The model fit (accuracy) to the observed data is evaluated using the Coefficient of Determination (CoD) as:

$$\text{CoD}=1-\frac{\sum_n (\hat{y}-y_{exp})2}{\sum_n(y_{exp}-avg(y_{exp}))2}=1-\frac{SSE}{\sum_n(y_{exp}-avg(y_{exp}))2} \quad (2.31)$$

Where n is the number of samples; ŷ is the estimated output of the model; y_{exp} is the observed value; $avg(y_{exp})$ is the average value of the corresponding observations (evaluated on the n samples) and SSE is the sum of square error

Xu et al. (2011a) examined the failure of CI pipes considering length, diameter and age as explanatory variables. The individual pipes were *aggregated* in two different ways. First, based on the diameter and then based on both the diameter and the installation year creating 9 and 501 groups respectively. In the first case, the age was calculated as the length-weighted age of all the pipes within each diameter group. The GP run resulted in a set of models and the selected models for the first (Eq. 2.32) and the second (Eq. 2.33) method are:

$$\text{B}=\frac{A(L+5.198)}{D-28.147} \quad (2.32)$$

$$\text{B}=(\frac{(40.47-\text{A})}{D-28.147}+1)\ \frac{AL}{D} \quad (2.33)$$

Where B is the predicted number of breaks, L is the total length, D is the diameter and A is the age

The Equation 2.33 was transformed to avoid negative values as follows:

$$B=\max\left(0, \left(\frac{(40.47-A)}{D-28.147}+1\right)\frac{AL}{D}\right) \quad (2.34)$$

The CoD for the Eq.2.32 was 0.994 and 0.951 for the model development and validation respectively. The very high values of CoD may be linked to the very small number (nine) of groups created implying that the GP model may suffer from over-fitting. In the second case where the number of groups increased (a few hundred instead of nine) and the CoD for the Eq.2.34 was significantly lower, 0.741 and 0.657 for the model development and validation phases respectively.

Xu et al. (2011b) *aggregated* the CI pipes into homogenous groups based on their diameter and age to avoid potential overfitting problems. Then, they partitioned the database into two clusters of those installed before and after the beginning of the monitoring period. Two distinctive GP models were developed, one for each cluster using the same explanatory variables (i.e. length, diameter and age) as Xu et al. 2011a. The implementation of the methodology resulted in a set of GP models for each cluster and the selected models for the pipes installed after (Eq. 2.35) and before (Eq. 2.36) the beginning of the monitoring period are:

$$BR_{GP}=\frac{14.928\ A^2\ L}{D^2} \quad (2.35)$$

$$BR_{GP}=\frac{44.446\ L\exp\left(-\frac{A}{70.246}\right)}{D} \quad (2.36)$$

Where BR is the predicted number of breaks, L is the total length, D is the diameter and A is the age

The CoD with the training dataset was 0.635 and 0.711 for the pipes installed after and before the beginning of the monitoring period respectively. The CoD for the validation dataset was very low, around 0.34 for both cases. This GP implementation attempted to address the over-fitting problem but failed to generate accurate predictions (as indicated by the very low values of CoD).

In both studies the explanatory variables are the length, the diameter and the age whereas environmental and operational variables have not been considered. Furthermore, GP lacks the capability to optimize coefficients efficiently and grows substantially in length very quickly (Davidson et al. 1999).

Evolutionary Polynomial Regression Models

Evolutionary Polynomial Regression (Table 2.5) is a data-driven method based on numerical and symbolic regression that can produce series of pseudo-polynomial models. EPR exploits both the power of evolutionary algorithms and numerical regression to develop polynomial models combining the independent variables together with the user-defined function. The user selects the generalised model structure, EPR employs a multi-objective search strategy to estimate unknown constant parameters of the assumed models using the least squares method (Giustolisi and Savic 2006). As a result, each single EPR run returns a number of polynomial models on a Pareto optimal front which is a trade-off between accuracy (fitness) and parsimony. The first criterion aims to maximise the model fit to the observed data (or minimise the model error) and the second (parsimony) aims to minimise the number of explanatory variables and/or polynomial terms in the model. The number of polynomial terms is a surrogate for the model parsimony criterion. Its role is to prevent over-fitting of the model to

data and thus endeavour to capture underlying general phenomena without replicating noise in data. The general expression of the EPR formula is given as (Giustolisi and Savic 2006):

$$\hat{y} = \sum_{j=1}^{m} F\left(X, f(X), a_j\right) + a_o \quad (2.37)$$

Where $\hat{y}$ is the estimated output of the system/process; a_j is a constant value; F is a function constructed by the process; X is the matrix of input variables; f is a function defined by the user; and *m* is the number of terms of the polynomial structured expression (bias a_0 excluded, if any)

Table 2.5 Summary of Evolutionary Polynomial Regression models

References	**Variables**
Berardi et al. (2008)	Length, diameter, age, number of properties supplied, number of pipes
Savic et al. (2009)	Length, diameter, age, number of properties supplied,
Xu et al. (2011b)	Length, diameter, age
Laucelli et al. (2014)	Average of mean daily temperature, minimum daily temperature, maximum daily temperature, maximum temperature increase, maximum temperature decrease, increase in temperature gradient, decrease in temperature gradient, freezing index, daily variation of temperature, snow cover

Berardi et al. (2008) *aggregated* the individual pipes based on the same age and diameter and applied the EPR method considering length, diameter, age, number of pipes and number of properties supplied as candidate explanatory variables.

The EPR run resulted in a set of models with varying number of explanatory variables and polynomial terms on the Pareto front and the selected one (Eq. 2.38) highlighted that pipe age (A), diameter (D) and pipe length (L) are the most important variables. The number of properties supplied did not substantially improve the predictions.

$$BR=\frac{0.084904\ A\ L}{D^{1.5}} \tag{2.38}$$

The derived *aggregated* model was used to estimate the burst rate (expressed in number of failures per year) at the individual pipe level due to the membership of pipe i to the class as follows:

$$\lambda_i=\frac{L_i}{L_{class}}\ \frac{BR_{class}}{T}=\frac{L_i}{L_{class}}\ \frac{a_1\{D_{class}^{\delta}\ L_{class}^{\gamma}\ P_{class}^{p}\ N_{class}^{\mu}\ A_{class}^{\alpha}\}}{T} \tag{2.39}$$

Where BR_{class} the number of bursts predicted by the *aggregated* model for the class the pipe *i* belongs to, L_{class} is the length of the class, L_i is the length of pipe i, T is the monitoring period, a_1 is the regression coefficient and δ,γ,p,μ,α the exponents selected in the EPR model

The EPR run returned mathematical equations that are ‘’transparent’’ and the user could select one of the models on the Pareto front considering both the accuracy and the parsimony criteria. The selected one-polynomial EPR model was effective in terms of regression performance (CoD=0.822). However, the proposed methodology for estimating the failure rate of individual pipes was not tested on a validation data set.

Savic et al. (2009) used a multi-objective genetic algorithm (MOGA) strategy within EPR based on the Pareto dominance criterion. The models obtained (Eq.

2.40-2.42) are ranked according to the number of terms. The proposed methodology was applied to 48 water quality zones within a UK WDN using length, diameter, age, and the number of properties supplied as candidate explanatory variables.

$$BR=a_i\ L^{1.5} \quad (2.40)$$

$$BR=a_i\ A^{1.5}L \quad (2.41)$$

$$BR=a_i\ \frac{A^{1.5}L}{D} \quad (2.42)$$

Where BR is the predicted number of breaks, L is the total length, D is the diameter and A is the age and a_i is a coefficient calculated separately for each water quality zone

The accuracy (in terms of CoD) varies significantly for the water quality zones. It ranges between negative values and 0.9339 for the first model, between 0.3410 and 0.9511 for the second model and between 0.3331 and 0.9833 for the third model. The inclusion of more explanatory variables increased the accuracy for most of the examined water quality zones. Savic et al. (2009 used the same explanatory variables as Berardi et al. (2008) excluding the number of properties supplied from their analysis as a non-influential factor.

Xu et al. (2011b) followed the methodology described in the previous subsection for data *aggregation* and variables selection to develop two distinctive EPR models. The implementation of EPR resulted in two sets of Pareto front models. The selected models for the pipes installed after (Eq. 2.43) and before (Eq. 2.44) the beginning of the monitoring period are:

$$BR_{EPR}=\frac{14.935\ A^2\ L}{D^2} \quad (2.43)$$

$$BR_{EPR}=\frac{157.4514\ L}{D\ A^{0.5}} \quad (2.44)$$

Where BR is the predicted number of breaks, L is the total length, D is the diameter and A is the age

The CoD was 0.635 and 0.711 for the pipes installed after and before the beginning of the monitoring period respectively. Both selected models showed a low accuracy. The comparison between the selected EPR and GP models shows that both approaches achieve the same accuracy. The only difference resides in the coefficient because in the EPR was obtained by means of a LS method, while in GP it was calculated through crossover and mutation.

The first EPR implementations (e.g. Berardi et al. 2008; Savic et al. 2009; Xu et al. 2011b) considered only physical variables (i.e. length, diameter and age). Laucelli et al. (2014) first considered environmental variables. They extended the work done by Rajani et al. (2012) and investigated the relationship between climate data and bursts of 150mm CI pipes for three non-overlapping time steps (lasting 5, 15 and 30 days) using average of mean daily temperature, minimum daily temperature, maximum daily temperature, maximum temperature increase, maximum temperature decrease, increase in temperature gradient, decrease in temperature gradient, freezing index, daily variation of temperature and snow cover as explanatory variables. The analysis was conducted separately for the warm and cold season with a different set of covariates. The two selected models for the cold and the warm season are:

$$BR=a_1\frac{FZI\ TDG^{0.5}}{ADD^2}+a_0 \quad (2.45)$$

$$BR=a_1\frac{MTD^2}{TRN^{0.5}ADD^2}+a_0 \quad (2.46)$$

Where BR is the number of failures per year for each season, FZI is the freezing index, TDG is the decrease in temperature gradient, ADD is the daily variation of the temperature, MTD is the maximum temperature decrease, TRN is the total rain and a_1, a_0 are constant coefficients which are calculated separately for the cold and warm season of each year of the monitoring period

The results indicated that the 30 days' time step provided the most accurate predictions for both seasons. However, the accuracy was high only for the cold season, the CoD was 0.76 instead of 0.45 for the warm season. Furthermore, the distinction between cold and warm period relies on specifying an arbitrary threshold and the distinction must be done in advance to select the appropriate mathematical relationship for pipe failure predictions.

2.4 Impacts Assessment

The impacts of pipe failure imply a cost which can be either direct, indirect or social (Rajani and Kleiner 2002). Direct cost includes the cost of lost water, the cost of breakage repair and the cost of direct damage to the property, the indirect the disruption in a commercial property, the accelerated deterioration of nearby infrastructures and the social cost the quality of life and the water quality (Makar and Rajani 2000). The indirect costs are more difficult to be quantified with a monetary value (Muhlbauer 2004; Pietrucha-Urbanik and Studziński 2018*).* Water quality problems (i.e. discolouration) caused by pressure and flow disturbances triggered by a pipe burst have been reported but the quantification of impacts is difficult and often only surrogate measures are used (Sadiq et al. 2005; Sadiq et al. 2006). Complexity also stems from the differing social situations and the varying vulnerability of affected customers (Bicik et al. 2009).

Michaud and Apostolakis (2006) proposed a scenario-based methodology for the ranking of the elements of a water-supply network according to their value to the network's owner. They employed a hierarchical value tree to aggregate impacts of pipe isolation using the Multi-Attribute Utility Theory. Several types of customers were considered by the authors. Their strategic assessment did not consider locations of isolation valves and neither used a hydraulic model to evaluate the full effect of segment isolation (e.g. low-pressure problems). They merely used graph theory to evaluate the impact of segment isolations.

Filion et al. (2007) proposed a stochastic design of WDS considering the impact of low and high-pressure failures in WDS. They quantified the consequences of a failure using expected annual damages sustained by residential, commercial, and industrial users.

Lindhe et al. (2009) carried out a probabilistic risk analysis using fault tree analysis on an integrated level. They evaluated the applicability of Customers Minutes Lost (CML) (proposed by Bakker et al. 2012) as a measure of impacts in two situations when no water was delivered (quantity failure) and when water was delivered but did not comply with water quality standards (quality failure). Both hard data and expert judgements were used to estimate probabilities of events and uncertainties in the estimates. Incorporation of expert judgements is facilitated by using the mean failure rate and mean downtime to model estimates of probabilities. The calculation of consequences included the duration of failure and the number of people affected. By multiplying the two attributes, the consequences are expressed in terms of CML. They did not consider the cases where the supply is partly satisfied but still in acceptable levels.

The fraction of delivered demand and the CML have been widely used for assessing the impacts on the level of service (e.g. Ostfeld et al. 2002; Ang and Jowitt 2006; Kapelan et al. 2006b; Giustolisi et al. 2008a; Jun and Guopin 2012). An exact quantification of failure impacts (particularly the water quality problems) is a highly subjective and complex problem (Rajani and Kleiner 2002; Sadiq et al. 2004c; Bicik et al. 2009).

2.5 Summary and Conclusions

The previous sections exhaustively described the developed methodologies for pipe failure prediction, impacts assessment and risk analysis.

Aging water pipes present problems including rehabilitation and renewal costs, decreasing hydraulic capacity, degradation of water quality, increasing customer complaints, and direct and indirect economic consequences due to service disruption. Hence, there is need to optimise the economic efficiency of the asset management in parallel with the satisfaction of the operational requirements. Proactive approaches using predictive analyses aim to achieve longer-term economic efficiency. Development of an accurate prediction model is important for the successful implementation of a proactive approach.

The pipe failure patterns present a high variability and therefore a more precise approach is required to accommodate it and improve the accuracy of predictions. Several methods have been developed for pipe failure prediction. Statistical models can cope with the lack of sufficient knowledge related to the complex mechanisms that lead to pipe failure but have some limitations such as requirements for specific assumptions (e.g. selection of probability distribution function) that should be substantiated by some knowledge of the phenomenon, which is not always available.

On the contrary, the data-driven methods can simulate complex non-linear relationships between inputs and outputs in the absence of physical consideration. The direct comparison between statistical and data-driven models (e.g. Achim et al. 2007; Asnaashari et al. 2013) has demonstrated their superiority in making predictions. Majority of the developed GP models have a low accuracy as indicated by the CoD values and the ones with high accuracy may suffer from overfitting due to the very small number of input data. EPR is similar to GP in terms of the class of results it generates (symbolic formulas), but it circumvents some of GP's shortcomings by integrating a genetic algorithm (Goldberg 1989; Rezania et al. 2011). Therefore, GP was deselected from the model development stage. Both ANN and EPR are powerful approaches and have demonstrated a high accuracy. However, the ANN require the optimum structure of the network (e.g. number of inputs, hidden layers, transfer functions) to be identified a priori (Giustolisi and Laucelli 2005; Rezania et al. 2011) which can be a time-consuming process and due to their black-box nature they do not provide a clear insight into the relationship between the inputs and the output in addition to data-fitting (Xu et al. 2013). On the contrary, each EPR run returns transparent equations that allow to gain an understanding into the relationship between the inputs and the output. The user can evaluate the models looking at different key aspects which encompass the prior knowledge of the phenomenon.

Models that can estimate the total number of failures in the network accurately can be used for economic analysis in long-term planning. Despite the accuracy obtained so far, there is still need for improvement, particularly for groups of pipes with a very high or very low failure rate. The inclusion of a clustering method can improve predictions by using a set of predictive models instead of a single-one which captures various failure patterns. Furthermore, a method with a high level

of accuracy assists in deriving the failure rate of individual pipes to point out the most prone to failure.

Most existing models estimate the number of failures or the average failure rate of the entire network or a group of pipes over a time-period (Selvakumar and Tafuri 2012; Scheidegger et al. 2013). For the following year(s) water utilities require more accurate and relatively simple models if possible (Chik et al. 2016). The failure frequency exhibits an inter-year variation which needs to be captured. The pipe deterioration rate is affected (among other) by time-dependent factors such as the weather conditions that are random or seasonal over time. The annual predictions facilitate the need for accurate annual predictions using both weather-related and physical factors.

There is empirical knowledge that in the UK most failures occur during the coldest months. Previously developed approaches (e.g. Rajani and Kleiner 2012; Laucelli et al. 2014) resulted in relationships with low accuracy for short-term predictions (i.e. a couple of days). The short-term predictions enable the water utility to adjust the daily allocation and planning of resources to accommodate possible increases in pipe failure, particularly for WDNs that are influenced by climate factors (Chik et al. 2018). The water utility is aided to meet the standards set by OFWAT to avoid customers dissatisfaction and compensations and comply with the guidelines for water loss.

WDNs are designed to satisfy the design flow and head at each demand node. However, in the case of a pipe failure, the water flow will change, and the original network will be transformed into a new one with higher internal energy losses which might make it impossible to deliver the desired flow rate at a minimum delivery pressure (Farmani et al. 2005). Hence, there is need to associate the satisfied demand with the nodal pressure changes to assess the performance of

the network in a realistic way. The magnitude and the scale of the impacts depend on many factors, amongst which, geographic location of pipe failure, the time of pipe failure and its duration and the topology of the WDN are some of the most important.

This thesis proposes a method to evaluate its impacts considering the ratio of unsupplied demand, the number of nodes with zero supply and the number of nodes with partial supply. The widely used ratio of unsupplied demand on its own is a rough indicator because for the same ratio the number of customers and the extent to which are affect can be different.

This thesis proposes a new combination of a grouped-based method for deriving the failure rate and an individual-pipe method for evaluating the impacts on the level of service. The conjunction of their outputs can identify the most critical pipes in the WDN.

Chapter 3. Pipe Failure Prediction

This study entails an approach for pipe failure prediction long-term, annually and short-term (Figure 3.1) considering a range of explanatory variables. The EPR method is employed for the long-term and the annual predictions while the ANN method for the short-term predictions.

As shown in Figure 3.1 the candidate explanatory variables of the long-term predictions are physical and operational factors while the environmental factors are the candidate explanatory variables of the annual and the short-term predictions.

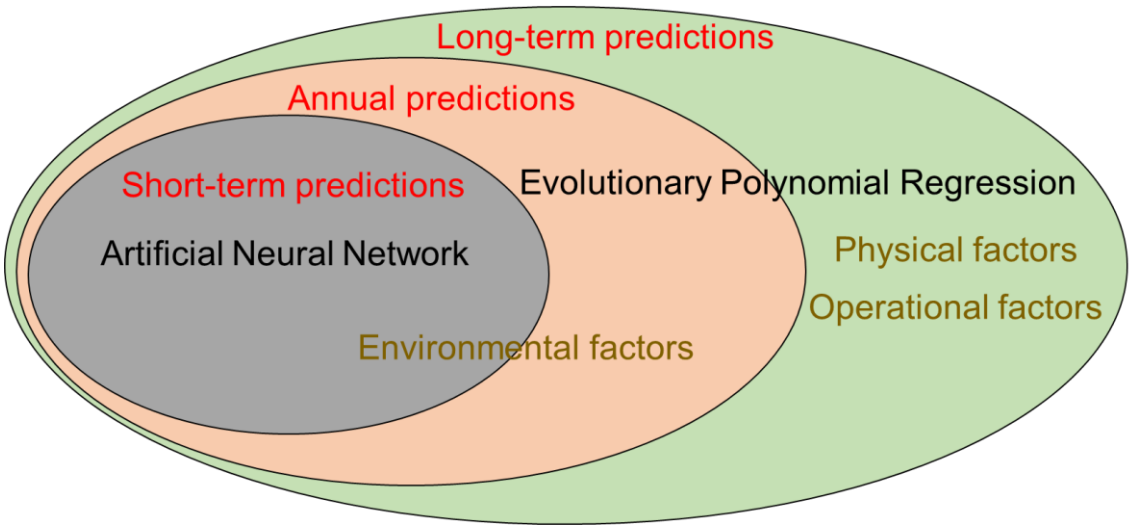

Figure 3.1. Outline of predictive models

Awareness about the need for long-term rehabilitation planning of the aging water infrastructure has risen globally (Herz, 1998; Burns et al. 1999; Kleiner and Rajani 1999; Engelhardt et al. 2000; Vanier 2001; Watson et al. 2004). The optimal management strategy for a WDN balances issues of water safety, reliability, quality, and quantity, while exploiting the full extent of the useful life of the pipes

and reducing long-term costs through proactive asset management (Kleiner and Rajani 2001). Pipe failure models are one of the key tools to support the M/R/R plans (Scheidegger et al. 2013; Francis et al. 2014; Scholten et al. 2014; Kabir et al. 2016). The long-term and the annual predictive models are developed to facilitate the need for effective asset investment.

All the UK water companies must follow the Guaranteed Standard Scheme (GSS) set out by OFWAT. If a company doesn't meet these standards, the affected customers may be able to claim compensation. If pressure falls below a specific threshold on two occasions, each occasion lasting more than one hour, within a 28-day period, the company must automatically make a GSS payment. The short-term predictions enable the water utility to adjust the daily allocation and planning of resources to accommodate possible increases in pipe failure.

It should be noted that the long-term and annual predictions are entirely separate from the short-term predictions since they are used for different purposes and they can’t be used in conjunction.

3.1 Methodology for Long-term Predictions

The long-term predictive models aim to predict the total number of pipe failures for the examined period and consists of the following steps:

1. Initially, the individual pipes are *aggregated* into homogenous groups using pipe descriptive variables and environmental factors. This is based on the assumption that pipes with similar specific physical properties such as material, diameter and age are expected to have the same breakage pattern (Kleiner and Rajani 2012). In addition to the pipe characteristics, soil type is used as an *aggregation* criterion based on the fact that the soil properties have been associated with the corrosion of the metallic pipes (Doyle et al. 2003; Sadiq et al.

2004a) which is a dominant factor contributing to their failure (Makar 2000; Rezaei et al. 2015). This was also confirmed by the preliminary analysis of asset and failure data for the case study used in this PhD research (Chapter 5.3). Each *aggregated* homogenous class of pipes takes a length and a number of failures equal to the total lengths and total number of failures for the individual pipes. Note that both failed and non-failed pipes are considered here. The original dataset containing individual pipes is converted to a new dataset containing homogenous groups of pipes based on diameter, age and soil type.

2. The created homogenous groups are split into 10 equal size folders using the cross-validation technique (Geisser 1975) for calibration and validation purposes.

3. The training dataset (i.e. 90% of the data) is partitioned into k clusters based on the age and the diameter. The clusters are created using the KMEANS function in MATLAB (® R2014b). The failure rate, which is the target variable, is excluded from the clustering stage to avoid a circular filling pitfall.

4. Develop k EPR models each associated with the training data of the relevant cluster. EPR-MOGA-XL vr.1 (Giustolisi and Savic 2009; Giustolisi et al. 2009) is employed to develop the EPR models. The test dataset (i.e. 10% of the data) is not used in the model construction phase, allowing the evaluation of model's ability to handle unseen data. The specific model structure considered here is (Giustolisi and Savic 2006):

$$\hat{y} = a_0 + \sum_{j=1}^{m} a_j \left((X_1)^{ES(j,1)} \dots \dots \dots . (X_k)^{ES(j,k)} \right) \quad (3.1)$$

Where $\hat{y}$ is the estimated output of the system/process, X_k is the kth explanatory variable, ES is the matrix of unknown exponents, a_j are the unknown polynomial coefficients (i.e. model parameters) and *m* is the maximum number of polynomial terms (in addition to the bias term a_0)

The candidate inputs (X_k) are set to be the total group length (L), the diameter (D) and the age (A), the average pressure (P_{ave}), the maximum pressure (P_{max}) and the output ($\hat{y}$) is set to be the number of failures (F).

For each homogenous group the pressure-related candidate explanatory variables are calculated as the ''average values'' of all the individual pipes.

The candidate exponents (ES) considered are: [-2, -1, -0.5, 0, 0.5, 1, 2]. The absolute values of the candidate exponents are used to describe potential square, linear or half-power relationship between the inputs and the output. The value 0 is chosen to deselect input candidates without influence on the output (Shahnazari et al. 2014), while the positive and negative values are considered to describe potential direct and inverse relationship between the inputs and the output of the model.

The maximum number of polynomial terms is chosen 2 (m=2) excluding the bias term to ensure the best fit without unnecessary complexity. Unnecessary complexity is defined as the addition of new terms that fit mostly random noise in the raw data rather than the underlying phenomenon (Savic et al. 2009). The Least Square (LS) method is constrained to search for positive polynomial coefficient values only ($\mathrm{i.e.}\, a_j>0$). The presence of negative coefficients polynomial coefficients ($\mathrm{i.e.}\, a_j<0$) may be required for a better description of the background noise in the data (Giustolisi et al. 2007).

5. Finally, the performance of the models is evaluated by using the test data. The Euclidian distance of input variables between the test data sample and the counterpart cluster centre values (known as centroids) is calculated to identify the suitable cluster for each test data. The corresponding EPR model associated with the relevant cluster is used to predict the number of pipe failures. By calculating the number of failures using the k EPR models for all test data samples,

performance indicators can be evaluated by using the predicted number of failures for the test dataset and the corresponding observations.

6. Various numbers of clusters (k) are tested until no further improvement is achieved for both the training and test datasets and too few groups are allocated in each cluster to identify the optimal number of clusters.

7. The aggregated EPR models are used to calculate the failure rate of individual pipes. It is assumed that all the individual pipes within the same homogenous group share the same failure rate and this is the average of the entire examined period.

8. Finally, the individual pipes are allocated in bands based on their failure rate using the Jenks natural breaks method (Jenks 1963) and the outputs are visualized using the ArcGIS mapping tool to illustrate the most prone to failure parts of the WDN (Vairavamoorthy et al. 2007).

3.1.1 Cross-validation Technique

Cross-validation is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set (Kohavi 1995). It involves partitioning the data into complementary subsets, performing the analysis on one subset (training set) and validating the analysis on the other subset (test set). The selection of training and test datasets has a significant impact on the results (Javadi et al. 2006). Cross validation consists of averaging several hold-out estimators of the model performance corresponding to different data splits (Geisser 1975) to correct for the optimistic nature of training error and derive a more accurate estimate of model prediction performance (Seni and Elder 2010). The advantage of the cross-validation technique used over repeated random sub-sampling is that all observations are used for both training and test,

and each observation is used for test exactly once (Gandhi et al. 2011). With the repeated random sub-sampling method some of the original data may be selected more than once in the test dataset whereas some other may not be selected at all. The cross-validation method used in this paper (Figure 3.2) is composed of the following steps:

- Divide the original data set into m folds

  For i=1,….,m:

- Train the model using all the data that do not belong to the i fold
- Test the model on fold i
- Estimate the average values of performance indicators

1st iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 2-10, Test folder: 1
2nd iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 1 & 3-10, Test folder: 2
3rd iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 1-2 & 4-10, Test folder: 3
4th iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 1-3 & 5-10, Test folder: 4
5th iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 1-4 & 6-10, Test folder: 5
6th iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 1-5 & 7-10, Test folder: 6
7th iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 1-6 & 8-10, Test folder: 7
8th iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 1-7 & 9-10, Test folder: 8
9th iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 1-8 & 10, Test folder: 9
10th iteration	1	2	3	4	5	6	7	8	9	10	Training folders: 1-9, Test folder: 10

Figure 3.2 10 folds cross-validation technique

The original data are randomly split into ten equal and mutually exclusive subsamples. Of the ten subsamples, nine subsamples (90% of the original data) are used as training data and a single subsample (10% of the original data) is retained for testing the model. Each of the ten subsamples is used exactly once as test data. The methodology is repeated ten times and the results of the training data from the ten iterations are averaged.

3.1.2 K-means Clustering Method

Data clustering is a technique that allows objects with similar characteristics to be grouped together for further processing (Pham et al. 2005; Kim and Keo 2015) and relies on the principle of simultaneously maximizing the intra-cluster similarity and minimizing the inter-cluster similarity (Wettschereck et al. 1997).

K-means (MacQueen 1967) is one of the simplest unsupervised learning algorithms that solve the clustering problem. It is a popular clustering method (Kanungo et al. 2002; Eghbali et al. 2017) due to its simplicity and efficiency (Sheng and Liu 2004). K-means assigns *n* data points into *k* clusters while minimizing an objective function of dissimilarity or distance (Jang et al. 1997). The k-means clustering moves objects between clusters until the objective function cannot be diminished further. In most cases the dissimilarity measure is chosen as the Euclidean distance. The objective function, based on the Euclidean distance, to be minimized can be defined by:

$$J=\sum_{j=1}^{k}\sum_{i=1}^{n}\left|x_i^{(j)}-c_j\right|^2 \quad (3.2)$$

Where $\left|x_i^{(j)}-c_j\right|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j is an indicator of the distance on the *n* data points from their respective cluster centres

The K-means algorithm is composed of the following steps (Redmond and Henegha 2007):

1. Initialise *k* cluster locations (c_1, c_2, …..c_k)
2. Assign each x_i to its nearest cluster centre c_k
3. Update each cluster centre c_k as the mean of all x_i that have been assigned as closest to it

4. Calculate the objective function *J*

5. If the value of *J* has converged, then return (c_1, c_2 …..c_k); else go back to Step 2

3.1.3 Jenks Natural Breaks Method

The Jenks natural breaks method (Jenks 1963) is designed to optimize the arrangement of a set of values into "natural" classes. This is achieved by minimizing each class's average deviation from the class mean (minimization of the variance within classes), while maximizing each class's deviation from the means of the other groups (maximization of variance between classes) (Jenks 1967). The Jenks method creates choropleth maps that have accurate representations of trends in the data. The advantage of the method is that it identifies real classes within the data (McMaster 1997) and is the default option of ArcGIS mapping tool with implementations in several fields (e.g. Holt et al. 2004; Brewers 2006; Kloog et al. 2008; Stefanidis and Stathis 2013).

It requires an iterative process which starts by dividing the data into predefined number of groups and is composed of the following steps:

- Calculate the sum of squared deviations between classes (SDBC).
- Calculate the sum of squared deviations from the array mean (SDAM).
- Subtract the SDBC from the SDAM (SDAM-SDBC). This equals the sum of the squared deviations from the class means (SDCM).
- After inspecting each of the SDBC, a decision is made to move one unit from the class with the largest SDBC toward the class with the lowest SDBC.
- New class deviations are then calculated, and the process is repeated until the sum of the within class deviations reaches a minimal value.

3.2 Methodology for annual predictions

This method is complimentary to the previous one which calculated the total number of failures for the entire examined period. However, this failure rate is not constant over this period (Figure 5.5). Therefore, there is a need to associate the annual number of pipe failures with factors that are dynamic (i.e. weather-related factors). The methodology for the annual predictions consists of the following steps:

1. The created homogenous groups (described in the above section) of pipes are allocated into the same clusters as the previous approach.

2. The first step in applying the EPR is the establishment of the inputs and the output. The 'dependent variable' is the annual number of failures (C_i) for each cluster on a yearly basis and the candidate 'explanatory variables' is a set of weather factors corresponding to this year (Eq. 3.3-3.7). The candidate explanatory variables are: average minimum air temperature (Eq. 3.3), average maximum air temperature (Eq. 3.4), average soil temperature (Eq. 3.5), freezing index (Eq. 3.6) and precipitation (Eq. 3.7).

$$AveT_{min}=\frac{\sum_{j=1}^{m} T_{minimum}^{j}}{m} \quad (3.3)$$

$$AveT_{max}=\frac{\sum_{j=1}^{m} T_{maximum}^{j}}{m} \quad (3.4)$$

$$\text{AveST}=\frac{\sum_{j=1}^{m} ST^{j}}{m} \quad (3.5)$$

$$\text{FI}=\sum_{j=1}^{m}(\theta - MinT^{j}) \quad (3.6)$$

$$\text{Precipitation}=\sum_{j=1}^{m} \text{Precipitation(k)} \quad (3.7)$$

Where *m* is the number of days in the time step *i*, $T_{minimum}^{j}$ is the minimum daily temperature of day *j*, $T_{maximum}^{j}$ is the maximum daily temperature of day *j*, ST^{j} is the daily soil temperature of day *j,* θ is the air temperature threshold

The freezing index (FI) (Eq. 3.6) is defined as the cumulative minimum daily temperature below a specified air temperature threshold and acts as a surrogate for the severity of extreme air temperatures within a time step (Kleiner and Rajani 2002). The cross-correlation function in MATLAB (® R2014b) is applied to measure the similarity between the candidate air temperature thresholds (ranging from -2^0C to 4^0C) and the number of failures. The temperature threshold with the highest similarity is selected for calculating the value of the FI. The process is repeated separately for each cluster.

The candidate exponents (ES) considered are: [-2, -1, -0.5, 0, 0.5, 1, 2]. The absolute values of the candidate exponents are used to describe potential square, linear or half-power relationship between the inputs and the output. The value 0 is chosen to deselect input candidates without influence on the output (Shahnazari et al. 2014), while the positive and negative values are considered to describe potential direct and inverse relationship between the inputs and the output of the model. The maximum number of polynomial terms is chosen 1 (m=1) excluding the bias term to ensure the best fit without unnecessary complexity (Savic et al. 2009).

3. The created dataset is split into two parts for calibration and validation purposes respectively. The last year of the monitoring period is used as test dataset while the remaining years are retained for model development. An individual EPR model is developed for each cluster associated with the relevant training data using EPR-MOGA-XL vr.1 (Giustolisi and Savic 2009, Giustolisi et al. 2009). The test dataset is not used in the model construction phase, allowing the evaluation of model's ability to handle unseen data.

4. The EPR models are selected with respect to the goodness of fit to the observed data and the possibility to describe the physical phenomenon. Their accuracy is assessed using the observed data.

5. The outputs of the EPR models are used to allocate the number of failures to the homogenous groups of pipes within each cluster proportionally to their length.

6. The predicted number of failures for each homogenous group is used to calculate the failure rate of individual pipes within this group. It is then assumed that all the individual pipes within a homogenous group share the same failure rate.

7. The outputs of the methodology are combined with the previous approach (long-term approach) to estimate the final failure rate.

8. Finally, the individual pipes are allocated in bands based on their failure rate using the Jenks natural breaks method and the outputs are visualized using the ArcGIS mapping tool.

3.3 Methodology for short-term predictions

The failures in a WDN require fast response from the operators. The water companies aim to respond as soon as possible after a burst is reported to minimize the amount of lost water and minimize the customer dissatisfaction. The response time depends on the factors such as the available resources. The short-term predictions facilitate the need for enhancing the daily allocation and planning of resources.

Contrary to the previous models and due to the limited number of failures in some clusters this approach is applied to the entire dataset to obtain models with a meaningful ‘goodness of fit’. This approach does not associate the failure

occurrence with specific pipes (e.g. pipe attributes and soil type) and is applicable merely for general operational purposes. The methodology consists of the following steps:

1. Define the inputs and the output of the binary model. The inputs of the ANN model are: the minimum air temperature, the maximum air temperature, the mean air temperature, the soil temperature and the FI while the targeted output of the model is 1 if there is at least a pipe failure the following days and 0 if not. The temperature variation can occur relatively quickly whereas the potential pipe failure because of that might take longer (Rajani and Kleiner 2012; Laucelli et al. 2014); therefore, the time step should be selected carefully. Various combinations of inputs and outputs (i.e. all the possible combinations from one up to seven days as input and from one up to seven days as output), are investigated to get the highest accuracy. The selection of failure/not failure as an output is case-specific and relies on the fact that for most of the days one or no failures are observed.

2. The original dataset (i.e. the entire network) is divided into 10 equal size folders using the cross-validation method and they are allocated as: 70% of the data (i.e. seven folders) for training, 20% of the data for validation (i.e. 2 folders) and the remaining 10% for test. The model is initially fit on a training dataset that is a set of examples used to fit the parameters (i.e. weights of connections between neurons) of the network (Ripley 2007; James et al. 2013). The validation dataset provides an unbiased evaluation of a model fit on the training dataset while tuning the model's hyperparameters (i.e. the number of hidden units in the neural network) (Ripley 2007; James et al. 2013). It is also used for regularization by early stopping: stop training when the error on the validation dataset increases, as it is a sign of overfitting to the training dataset (Prechelt 1998; Asnaashari et

al. 2013). The test sample provides an unbiased evaluation of the final model fit. The test (holdout) dataset is not used in the training phase.

3. The actual output of the model is not an integer number; therefore, the optimal threshold for converging to 1 (failure) or 0 (non-failure) must be identified. The selection of the optimal threshold entails three steps:

3a. Initially is defined a set of candidate thresholds covering the entire range between the model's minimum and maximum responses for the test data. Then, the model's actual outputs are rounded for all the values of candidate thresholds. The threshold is a cut-off value, any probability greater than this threshold are considered as failure and the rest are considered as no failure.

3b. The True Positive Rate (TPR) and the False Positive Rate (FPR) are calculated for all the candidate thresholds. This iterative process provides a set of TPR/FPR pairs which are used to plot the Receiver Operating Characteristic (ROC) curve (Figure 3.3).

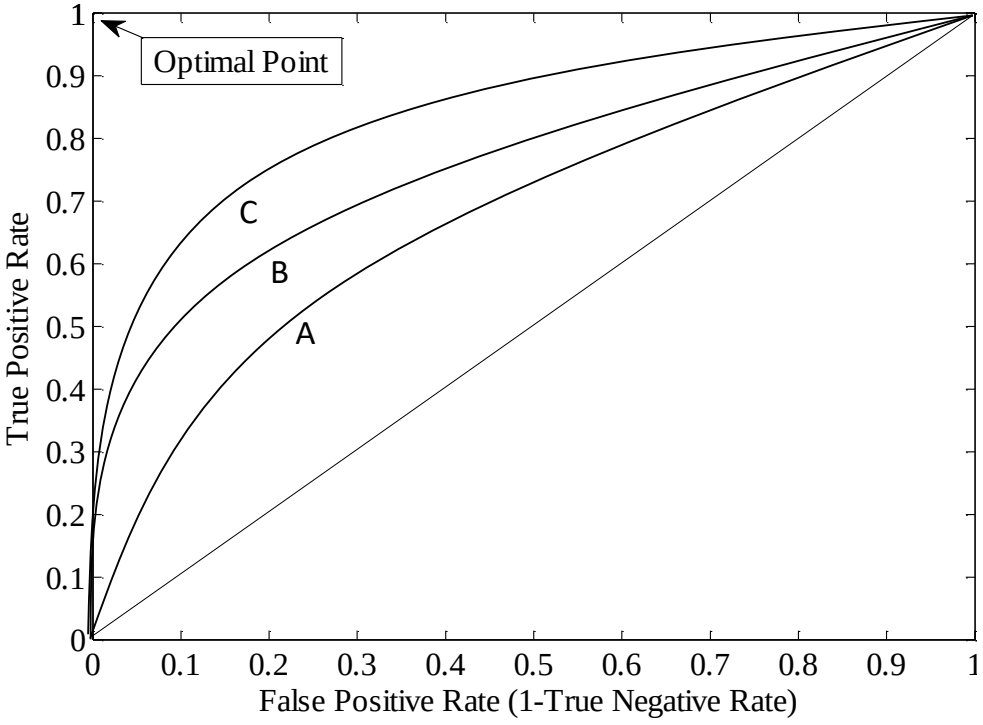

Figure 3.3: General form of Receiver Operating Curve

Each point on the ROC plot represents a specific TPR/FPR pair. A model with perfect discrimination has a ROC curve that passes through the upper left corner (optimal point) (Zweig and Campbell 1993). On the contrary, the closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the model is. Therefore, curve C has a higher accuracy than curves A and B, while curve B is more accurate compared to curve A. The ROC curve provides an overall representation of the accuracy and the model is good if it can discern the days with pipe breaks.

ROC graphs have recently gained attention in machine learning and data mining research in several fields including the water sector as well (e.g. Debón et al. 2010; Romano et al. 2014; Chik et al. 2016 Mounce et al. 2017; Winkler et al. 2018).

3c. The Euclidian distance (distance between each point on the curve and the optimal point) is calculated as follows:

Euclidian distance=$\sqrt{(1-TPR)^2+(FPR)^2}$ (3.8)

The threshold with the minimum Euclidian distance is selected since it simultaneously minimizes the false positive rate and maximizes the true positive rate.

4. Due to the black-box nature of the ANN model, the relative significance of each input variables has to be estimated. The influence of the input variables on the model's response is assessed as follows (Duncan et al. 2013):

W_{io}=W_1*W_2 (3.9)

Where: W_{io}= input-to-output influence vector; W_1= ANN hidden layer weight matrix; W_2= ANN output layer weight vector. Thus W_{io} has dimensions of N_{in}*N_{out} where N_{in} is the number of inputs and N_{out}=1 is the number of output neurons

3.4 Model performance assessment

The performance indicators used for the EPR regression models are the Coefficient of Determination (R^2) which is a measure for correlation between predictions and observations and Root Mean Square Error (RMSE) which is a measure for error predictions. Their equations are (Moriasi et al. 2007):

$$R^2=\frac{(\sum_{i=1}^{n}(y_{p,i}-\bar{y}_p)(y_{o,i}-\bar{y}_o))^2}{\sum_{i=1}^{n}(y_{p,i}-\bar{y}_p)^2\sum_{i=1}^{n}(y_{o,i}-\bar{y}_o)^2} \quad (3.10)$$

$$\text{RMSE}=\sqrt{\frac{\sum_{i=1}^{n}(y_{p,i}-y_{o,i})^2}{n}} \quad (3.11)$$

Where $y_{p,i}$ = prediction value for test sample *i*; $y_{o,i}$ = measurement value for test sample *i*, $\bar{y}_p$ = mean value of predictions, $\bar{y}_o$= mean value of measurements and *n* = the number of test data samples

The performance of the ANN model is assessed using the TPR (Eq. 3.12), the TNR (Eq. 3.13) and the Area under Curve (AUC). TPR measures the proportion of correctly identified positives while TNR measures the proportion of correctly identified negatives. They are expressed as follows (Kohavi and Provost 1998):

$$\text{True Positive Rate}=\frac{True\ Positives}{True\ positives+False\ negatives} \quad (3.12)$$

$$\text{True Negative Rate}=\frac{True\ Negatives}{True\ negatives+False\ positives} \quad (3.13)$$

The AUC calculated is a measurement of the performance of the model and larger values indicate a better overall performance (Kumar and Indrayan 2011). An AUC of 1 represents a perfect model whereas an AUC of 0.5 represents a worthless model. Equal AUCs of two tests indicate similar overall performance but the curves are not necessarily identical (Kumar and Indrayan 2011).

Chapter 4. Impacts Assessment

4.1 Introduction

The overall objective of a WDN is always to provide consumers with acceptable level of supply under a range of operating conditions (Atkinson et al. 2014; Large et al. 2015). However, certain conditions (e.g. pipe failure) can cause pressure and flow disturbances resulting in water quality issues, decline or interruption in system performance, loss of business, and costs associated with emergency response (Sadiq et al. 2006). Moreover, other nearby infrastructures such as pavement, road, storm water, sewer may be affected (Makar and Rajani 2000).

Asset management practices are used to identify investment strategies that avoid premature replacement of pipes while minimizing water main breaks, interruptions in service and costs of damage (Wood and Lence 2009). Localized repair, rehabilitation and replacement decisions enhance the condition of water mains to deliver acceptable level of service in terms of water demand and quality requirements (Engelhardt et al. 2003; Ammar et al. 2012) since it is impractical and unrealistic to replace all the aging pipes simultaneously due to budget limitations (Kleiner et al. 2001; Kim et al. 2004). An efficient maintenance program should identify the most vulnerable pipes whose failure could incur significant impacts (Jun et al. 2008; Kabir et al. 2015a). A break can cause water outage for all downstream consumers while a local pipe break may cause a water shortage for only one household (Wang and Chen 2015).

The pipe failure can result in pressure disturbances. Hence, the next section (4.2) presents the background of the pressure-driven analysis and the main advantages and drawbacks of each methodology.

It is followed by section 4.3 which entails the methodology for simulating the pressure deficient conditions and for quantifying the impacts of pipe failure on the level of service.

4.2 Pressure-driven Analysis

Pressure is a key factor in operating WDNs and should be carefully managed (Ghorbanian et al. 2016). To ensure safe and reliable delivery of water across a WDN, system pressure should generally be maintained between minimum and maximum acceptable levels (Ghorbanian et al. 2016). There are two types of WDN analysis, the demand-driven analysis (DDA) and the pressure-driven analysis (PDA).

The DDA hydraulic solvers are based on the well-known mass and energy balance equations that are used to compute pipe flows and nodal pressures in the network. These methods work well under normal flow conditions in which sufficient pressures are available in the network and that the available discharge in demand nodes is always equal to the required discharge (Reddy and Elango 1989; Tanyimboh and Templeman 2010; Bicik et al. 2011; Shirzad et al. 2012; Mahmoud et al. 2017). However, in case of pipe failure pressure can fall substantially and the DDA solvers can't always deliver realistic predictions of pressures and flows, because it is not always possible to deliver all desired demands under these circumstances (Gupta and Bhave 1996; Tanyimboh et al. 2001; Baek et al. 2010; Jun and Guoping 2012; Yoo et al. 2012). As a result, DDA solvers may produce unrealistically low, sometimes negative nodal pressures that are, in some cases, physically impossible (Tabesh et al. 2002; Kapelan et al. 2006a; Tanyimboh and Templeman 2010; Romano et al. 2014; Sayyed et al. 2015).

In the PDA the demand changes as the nodal pressure changes and the performance of a WDN can be assessed in a more realistic way (Chandapillai 1991; Fujiwara and Ganesharajah 1993; Kalungi and Tanyimboh 2003). The PDA methodologies can be broadly divided into two groups (Jun and Guoping 2012; Gorev and Kodzhespirova 2013; Sayyed et al. 2015).

The first group methods use a specific relationship between the nodal head and the flow (NHFR) (e.g. Bhave 1981; Germanopoulos 1985; Wagner et al. 1988; Fujiwara and Ganesharajah 1993; Gupta and Bhave 1994; Tanyimboh and Templeman 2010). During simulation, NHFR at different nodes must be satisfied along with node flow continuity and conservation of energy equations (Sayyed et al. 2014). One way to simulate pressure-deficient networks in EPANET (Rossman 2000) is to modify the source code of EPANET to get a direct solution that satisfies the NHFRs which can be a difficult task (Sayyed et al. 2014). Another way to solve this problem is to iteratively use EPANET with the node head-flow relationships satisfied externally in each iteration (Sayyed et al. 2015). This method is applicable for small WDNs but often time consuming and cumbersome especially for large systems (Babu and Mohan 2012). Most NHFR consist of three separate functions embedded into the governing system of equations for obtaining zero, partial and full nodal demands which causes an absence of continuity in their function derivatives at the transitions between zero and partial nodal flow and between partial and full demand satisfaction (Siew and Tanyimboh 2012).

Bhave (1981) categorised the outflow at a demand node as fully satisfactory if the head was not less than the head required at that node or zero if the head at that node was less than the elevation of the node. All other nodes were modelled as a ground level tanks to determine their outflows. Germanopolous (1985) used

a substitution relationship incorporated directly into the system of equations. This approach requires a smooth, continuously differentiable function of the head-outflow relationship if it is to work properly (Ackley et al. 2001). The NHFR proposed by Wagner et al (1988) (Eq. 4.1) is one of the most widely accepted (Gupta et al. 2013). It considers all the three operational modes in the distribution network: the normal mode (adequate flow), the deficient mode (partial flow) and the failed mode (no flow). The nodal flows for each corresponding mode are calculated based on the expressions given below:

$$q_i(t) = \begin{cases} d_i(t) & \text{if } P_i(t) > P_{des} \\ d_i(t)\left(\frac{P_i(t)-P_{min}}{P_{des}-P_{min}}\right)^{0.5} & \text{if } P_{min} < P_i(t) < P_{des} \\ 0, & \text{if } P_{min} > P_i(t) \end{cases} \quad \text{(Eq. 4.1)}$$

Where $q_i(t)$ is the estimated demand of node i at time t, $d_i(t)$ is the desired nodal demand at node i at time t when there is no pipe failure, $P_i(t)$ is the actual pressure at node i at time t, P_{min} is the minimum allowed pressure at node i, P_{des} is the desired pressure at junction i at time t

Experimental data from Shirzad et al. (2012) and Walski et al. (2017) justify the validity of this function to model nodal demands under critical pressures. The main difference between Wagner et al. (1988) and Fujiwara and Ganesharajah (1993) is that in Eq. 4.1 for lower heads, available outflow increases sharply, while, in the later the sharp increase of outflow relative to its previous state happens near desired pressure head. Gupta and Bhave (1996) made a comparison of various formulae that describe the pressure dependency of nodal consumption and they concluded that the following parabolic relationship (Eq. 4.2) provided by Chandapillai (1991) was sufficiently accurate.

$$H_j = H_j^{min} + R_j (Q_j^{avl})^{n_j} \quad \text{(Eq 4.2)}$$

Where H_j is the available head at node j, H_j^{min} is the minimum required head at node j; (i.e. the value below which outflow is assumed to be zero), R_j is a resistance constant and Q_j^{avl} is the available outflow at node j

The exponent n_j is both node and network specific and often varies between 1.5 and 2 (Gupta and Bhave 1996). Thus, the NHFR is transformed as follows:

$$q_i(t) \begin{cases} d_i(t) & \text{if } P_i(t) > P_{des} \\ = d_i(t)\left(\frac{P_i(t) - P_{min}}{P_{des} - P_{min}}\right)^{1/n_j} & \text{if } P_{min} < P_i(t) < P_{des} \\ 0, & \text{if } P_{min} > P_i(t) \end{cases} \quad \text{(Eq. 4.3)}$$

Where $\mathrm{q_i(t)}$ is the estimated demand of node i at time t, $\mathrm{d_i(t)}$ is the desired nodal demand at node i at time t when there is no pipe failure, $\mathrm{P_i(t)}$ is the actual pressure at node i at time t, $\mathrm{P_{min}}$ is the minimum allowed pressure at node i, $\mathrm{P_{des}}$ is the desired pressure at junction i at time t

The second group comprises of methods that determine nodal outflows implicitly without the need to introduce a NHFR function by adding to demand nodes a series of elements such as flow control valves (FCV), check valves (CV), pressure reducing valves (PRV), general purpose valves (GPV), throttle control valves (TCV), emitters (E) or artificial reservoirs (R) (e.g. Ozger 2003; Todini 2003; Ang and Jowitt 2006; Babu and Mohan 2012; Gorev and Kodzhespirova 2013; Sayyed et al. 2015; Mahmoud et al. 2017; Pacchin et al. 2017; Paez et al. 2018). This group of approaches does not require any parameters calibration (i.e. describing the pressure-demand relationship) unlike the first group (Bicik 2010).

There are two different methodologies to incorporate the artificial elements into the PDA analysis (Paez et al. 2018). They can either be progressively added and/or removed in each iteration (e.g. Ang and Jowitt 2006) or they can be assigned to all the demand nodes and thus the network's topology does not need to be modified iteratively (e.g. Paez et al. 2018).

Ozger (2003) and Ang and Jowitt (2006) connected (and removed when necessary) the reservoir to pressure-deficient nodes to calculate the actual flows delivered, followed by a DDA run in an iterative manner until convergence is achieved. The drawback of the method is that requires multiple runs of the EPANET network solver until a condition is reached where there are no nodes where water is withdrawn under negative pressure resulting in high computational costs (Rossman 2007; Wu 2007). The need to add and remove reservoirs at various stages of the iterative methodology makes its implementation difficult in large networks especially under extended period simulation (EPS) analysis (Wu 2007; Wu et al. 2009; Paez et al. 2018). The network topology modifications at each time step to identify the correct pressure-deficient and pressured deficient nodes lead to very slow convergence (Wu 2007; Wu et al. 2009; Kovalenko et al. 2014). Babu and Mohan (2012) extended the method proposed by Ang and Jowitt (2006) to carry out pressure-deficient network modelling in a single execution of the unmodified EPANET 2. They used a reservoir, a pipe with negligible resistance along with artificial FCV to prevent surplus flow into artificially added reservoirs and restrict the negative pressure in the network. Despite the smaller number of iterations compared to previous approaches, the algorithm does not model the transition between zero and full flow at a demand node satisfactorily and (Gorev and Kodzhespirova 2013). Gorev and Kodzhespirova (2013) tried to overcome this weakness introducing a series of artificial elements: a FCV, a pipe

(with a suitable resistance coefficient) with a CV and a reservoir at each demand node. Although the results were obtained with a single hydraulic run, the approach only supported the specific, parabolic type of NHFR (Wagner et al. 1988). Sayyed et al. (2014; 2015) replaced the reservoir and pipe with a suitably chosen flow emitter to reflect the properties of each node in the network. Pacchin et al. (2017) proposed adding a sequence of devices composed of a GPV, a fictitious junction, a reach with a CV without minor losses and an artificial reservoir at the demand nodes. The proposed method differs from other methods previously proposed in that it uses a GPV which allows the user to define the relationship between the supplied demand and available pressure, making the sequence of elements capable of representing different relationships among these variables. Paez et al. (2018) added in order, a FCV, a dummy junction, a TCV, another dummy junction, a CV and an artificial reservoir to each demand node of the network and tested it in two benchmark and a real complex network. The results showed that the method can simulate the network with pressure driven demands in EPS without modifying the EPANET2 source code or using its programmer's toolkit. The computational time was kept within acceptable range for most cases.

4.3 Methodology for impacts assessment

The exact quantification of impacts requires the knowledge of location, timing, and duration of failure and the topology of the network (Bicik 2010; Grigg 2013). However, the time when the failure occurs, and the duration of service disruption are not predictable (Vamvakeridou-Lyroudia et al. 2010; Grigg 2013; Berardi et al. 2014; Shuang et al. 2017). The proposed methodology consists of the following steps:

1. The first and paramount step is the simulation of pipe failure. Based on the conclusion derived from the previous section the pipe failure is simulated by adding a series of artificial elements They are added on a selective basis, i.e., only to pressure-deficient nodes with demands (i.e. demand nodes with available pressure less than the desired value). These nodes are identified by running the DDA-type hydraulic solver (i.e. EPANET) once before the PDA simulation. The examined network (Chapter 5.1) is a large real-life network and a single failure of a distribution pipe (not a transmission pipe which are typically between water treatment works and service reservoirs or between service reservoirs) is likely to cause pressure deficient conditions in a small part of it, and hence, the artificial elements are not added to all the demand nodes.

The sequence of artificial elements added to all the demand nodes with pressure-deficient conditions is (as suggested by Mahmoud et al. 2017) (Figure 4.1): a Check Valve, an internal dummy Node, a Flow Control Valve and an Emitter.

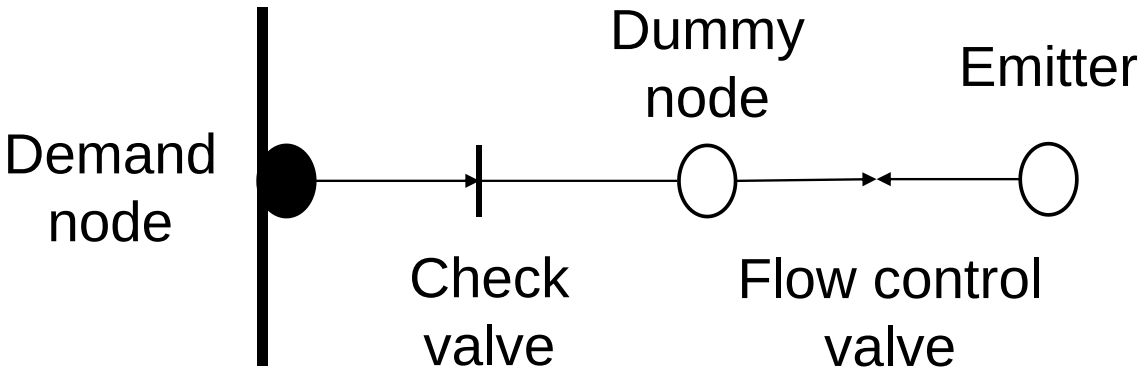

Figure 4.1 Artificial elements connected to a demand node

The role of a CV is to prevent flow reversal and is the first added artificial element on the demand node. The parameters of the CV are set to produce negligible head losses when water is flowing in the right direction (i.e. short length and large

diameter). The downstream dummy node is used just to connect the CV with the TCV since it is an EPANET requirement (Mahmoud et al. 2017; Paez et al. 2018). The role of the FCV is to ensure that the delivered flow does not exceed the demand at the node. Finally, the emitter is used to represent pressure-dependent demand delivery. The small length, large diameter, and large Hazen-Williams coefficient ensure that all additional elements do not introduce (significant) head loss between demand node and the emitter (Mahmoud et al. 2017). The parameter settings of the added elements are specified in Table 4.1.

Table 4.1 Parameters for artificial elements

Element	**Parameter**
Demand node	Demand ←Required
	Length ←small (i.e. 0.01m)
Check Valve	Diameter ←large (i.e. 1000mm)
	Setting ←maximum flow
Dummy Node	Elevation as the demand node
Flow Control Valve	Diameter ←large (i.e. 1000mm)
	Demand ←Required
Emitter	Emitter coefficient (C_d) ←Eq.4.4

The delivered flow to deficient nodes (*j*) is calculated as (Eq. 4.4):

$$Q_j^{avl} = \begin{cases} 0 & \text{if } H_j^{avl} < H_j^{min} \\ C_d (H_j^{avl} - H_j^{min})^{\gamma} & \text{if } H_j^{req} > H_j^{avl} \geq H_j^{min} \\ Q_j^{req} & \text{if } H_j^{avl} \geq H_j^{req} \end{cases} \quad \text{(Eq. 4.4)}$$

Where C_d is the estimated emitter coefficient as: $C_d = Q_j^{req}/(H_j^{req} - H_j^{min})$, and γ is the emitter coefficient estimated as γ=1/n_j

The values of both variables depend on the properties of each node that is defined in terms of empirical exponent coefficient n_j (value in the range between 0.5 and 2.5), and the characteristic heads (H_j^{req}, H_j^{min}).

2. The WDN's demand and pressure exhibit diurnal fluctuations as shown in Figure 4.2. The pressure in a WDN is minimum when the flow is maximum and coincides with the peak demand, whereas is maximum when demand is minimum, normally at night-time when most consumers are asleep and most industries do not operate (Jacobs and Strijdom 2009; Beal and Stewart 2013; Wang and Chen 2015).

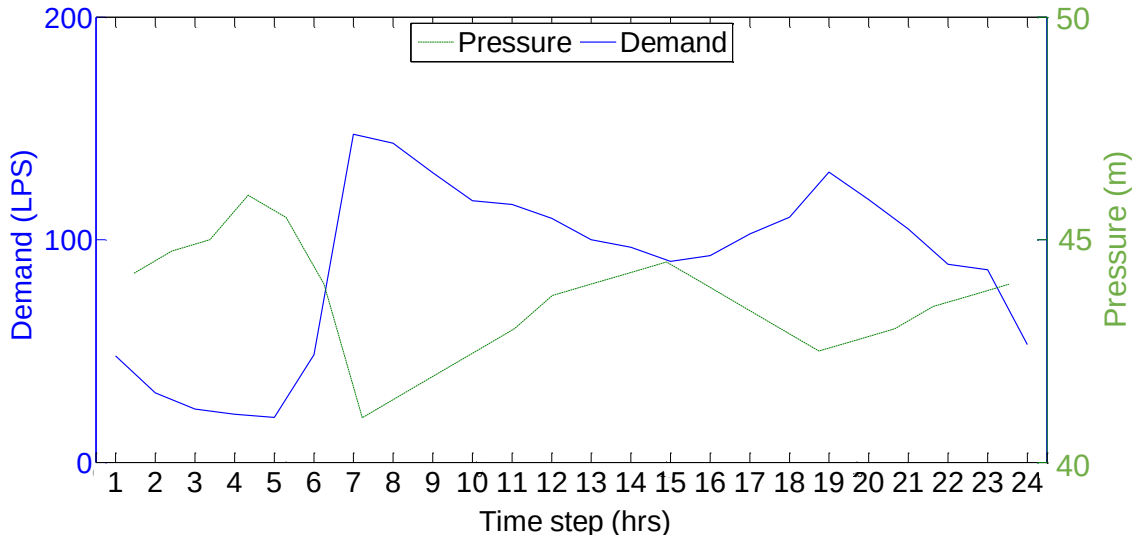

Figure 4.2 Diurnal variation of demand and pressure

The exact time of pipe failure can't be predicted and hence is assumed that the failure occurs when: 1) there is a peak in pressure (i.e. 4.00am) and 2) a peak in demand (i.e. 7.00am) as shown in Figure 4.2.

Failure duration is the time period from the start of the failure event to the completion of repair and is determined by unawareness, awareness, location, isolation and repair periods (Mounce and Boxall 2010; Bakker et al. 2012). During this time the network experiences two different phases: failure and isolation (Mansoor 2007). During the failure phase there will be a free flow of water from the crack on the pipe and the network will experience a drop-in pressure (Germanopoulos et al. 1986; Jowitt and Xu 1993). This dynamic situation lasts until the failed pipe is isolated from the network. After the isolation the networks starts to recover as the water flow is stopped, but still there will be nodes with inadequate pressure since the network operates at reduced mode (Jowitt and Xu 1993). This situation will remain until the repair is completed.

The pipe failure is examined in this thesis for an extended simulation period which is advisable in face of a single pipe burst (Berardi et al. 2014). The EPS is essentially a sequence of steady state simulations of the system whose boundary conditions are updated, sometimes according the last realization of system state, to reflect changes in nodal demands, tank levels, pump operations, etc. (Giustolisi et al. 2008). The diurnal demand variation in nodes, the water level in storage tanks, and the valve/pump control settings are considered over this predefined simulation period. by changing the parameters of the connected FCVs and emitters according to the current values of the desired demands in deficient nodes.

3. The performance indicators used for assessing the impacts of pipe failure on the level of service are:

- Ratio of unsupplied demand
- Percentage of nodes with zero satisfied demand
- Percentage of nodes with partly satisfied demand

Low pressure, although not being as severe as a complete interruption to water supply, causes inconvenience to customers and affects pressure sensitive water consumption.

The ratio of unsupplied demand is on its own is a poor indicator since for the ratio the number of nodes and extend to which they are affected can be significantly different. It is calculated as:

$$D=\frac{\sum_1^t \sum_1^i (d_i(t)-q_i(t))}{\sum_1^t \sum_1^i (d_i(t))} \quad \text{(Eq. 4.5)}$$

Where $q_i(t)$ is the estimated demand of node i at time t, $d_i(t)$ is the actual nodal demand at node i at time t when there is no pipe failure and t is the duration of the simulation period

The desired pressure threshold was set to 15m. WDN will operate normal at the pressure values above this. The minimum pressure was set to 0m. The threshold is derived for the examined case study based on the outputs of the calibrated EPANET model. Also, in the UK, OFWAT requires low pressure incidents (i.e. drops of pressure below 15m of head at water main) to be reported (i.e. as part of the DG2 PI) by every water utility.

5.1 General Analysis

The proposed methodology was demonstrated in a case study located in part of a WDN of a UK city. The WDN consists of cast iron (CI), ductile iron (DI), asbestos cement (AC), Polyvinyl chloride (PVC) and polyethylene (PE) pipes. The total length of the area is 833.10 (km) and the percentage of each material is shown in Figure 5.1.

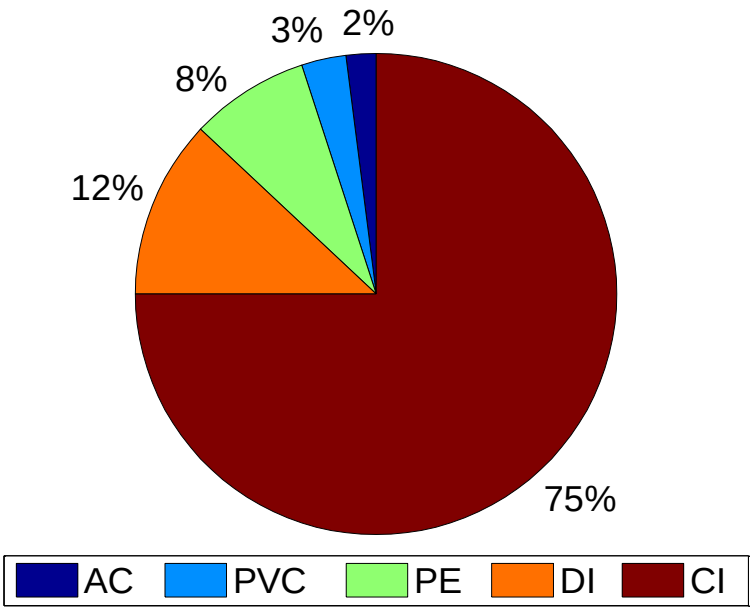

Figure 5.1 Percentage of different pipe materials in the network

The examined WDN consists of 33372 junctions, 8 reservoirs, 2 tanks, 27139 pipes, 10 pumps and 8037 valves. The information has been taken from an EPANET hydraulic model. The total number of customers and the number of customers per node that are fed by the network is not known.

The WDN entails pipes with the following diameters and installation years as shown in Figures 5.2 and 5.3 respectively. Most of the pipes are up to 150mm (mainly 100mm) while the installation year exhibits a significant variation.

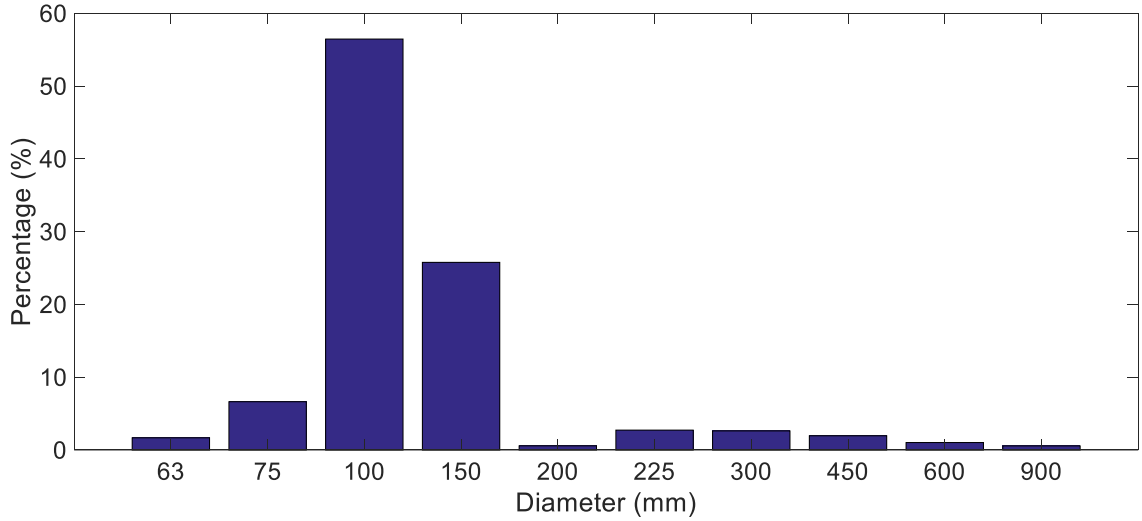

Figure 5.2 Percentage of pipes based on the diameter (mm)

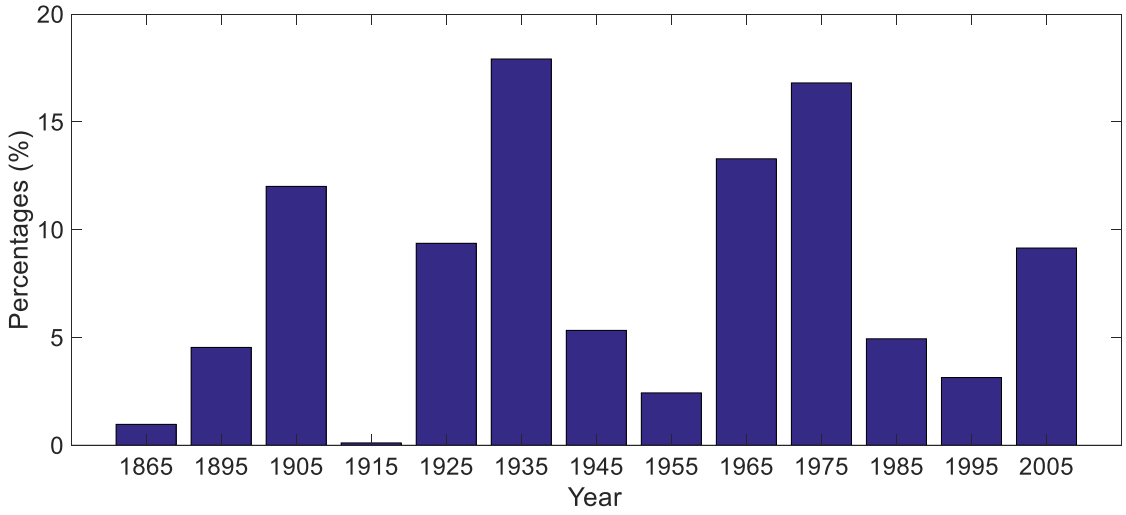

Figure 5.3 Percentage of pipes based on the installation year

The examined period lasts for 11 years, between 2003 and 2013 and the total number of failures for is 1810. Table 5.1 shows the average failure rate during this period (expressed in number of failures/km/year) for all the materials.

Table 5.1 Pipe Failure rate by the material type

Pipe material	Failure rate
AC	0.126
CI	0.203
DI	0.073
PE	0.037
PVC	0.131

Based on the comparison of the failure rate and the fact that CI pipes constitute 75% of the WDN's total length, it was decided only CI pipes to be included in the analysis. The CI pipes installed after the beginning of the monitoring period and with diameter greater than 300mm have been excluded from the analysis.

Most of the examined CI pipes (94.23%) have not experienced any failures, while 4.41% of them have failed once and 1.36% of the pipes have failed more than once during the examined period. Table 5.2 and Figure 5.4 show the main features of the CI pipes considered in the analysis.

Table 5.2 The main features of the CI pipes

Feature	Value/range
Installation year	1895-1995
Diameter range	75-300 mm
Total length	607 km
Number of pipes	18872
Number of failed pipes	1089
Number of failures	1369

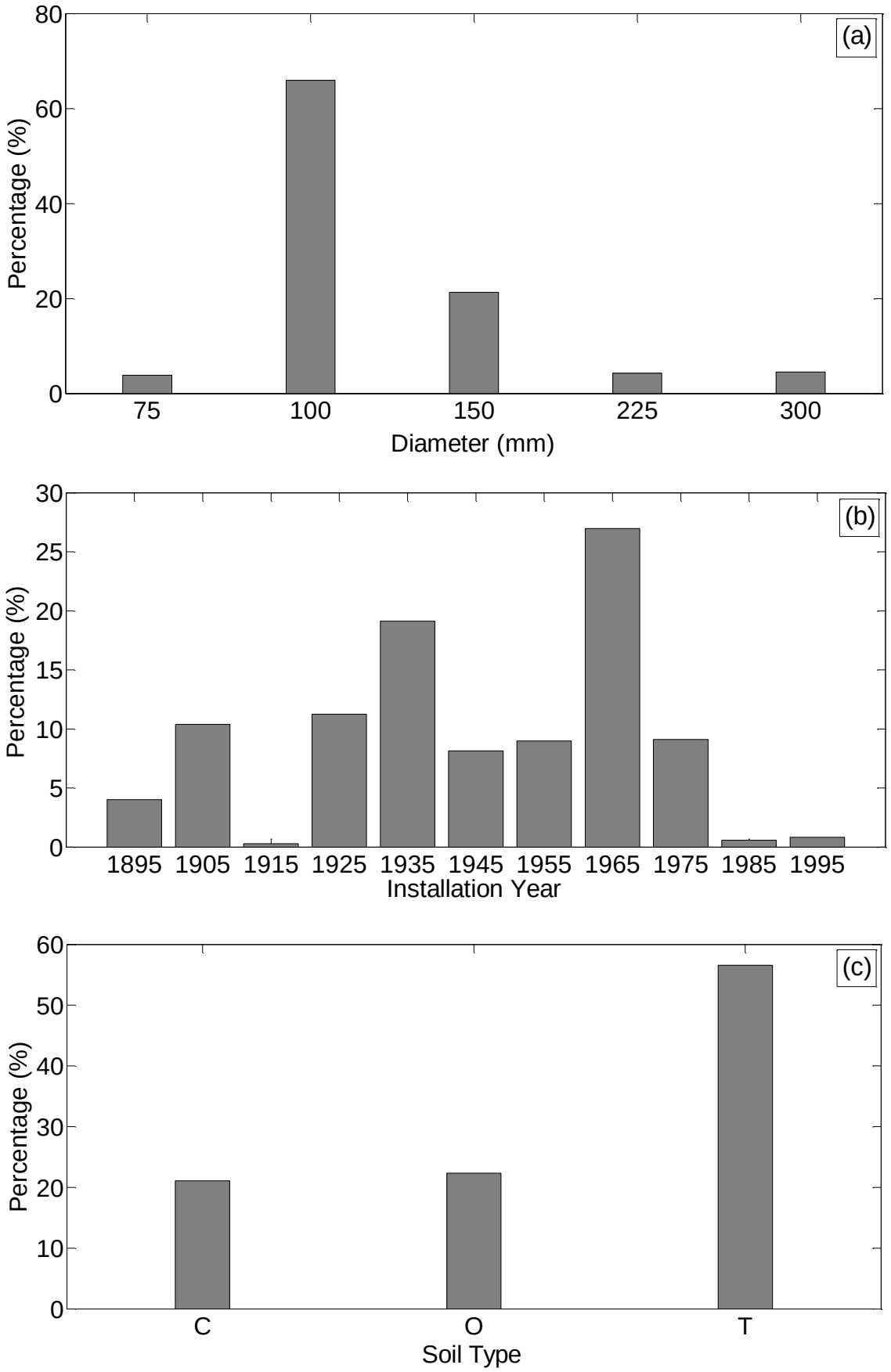

Figure 5.4 Percentage of CI pipes based on (a) diameter (mm), (b) age and (c) soil type

As shown in Figure 5.4 a big portion of the examined CI pipes is 100mm diameter while the installation years varies significantly for different groups. Some of the pipes have been laid in the ground for more than a century whereas some others were installed relatively recently. The pipes are installed in three different soil types (C, O and T respectively). Detailed information about the soil type are not available, only their initials are known.

The monitoring period lasts for 11 years (2003-2013), i.e. 4018 days. Most of the days (73.25%) there are not any failures, on 21% of the days only one failure occurred and more than one occurred on 5.75% of the days. The failures are not evenly distributed over the monitoring period. The preliminary analysis showed an intense intra-year and inter-year variation (Figures 5.5 and 5.6).

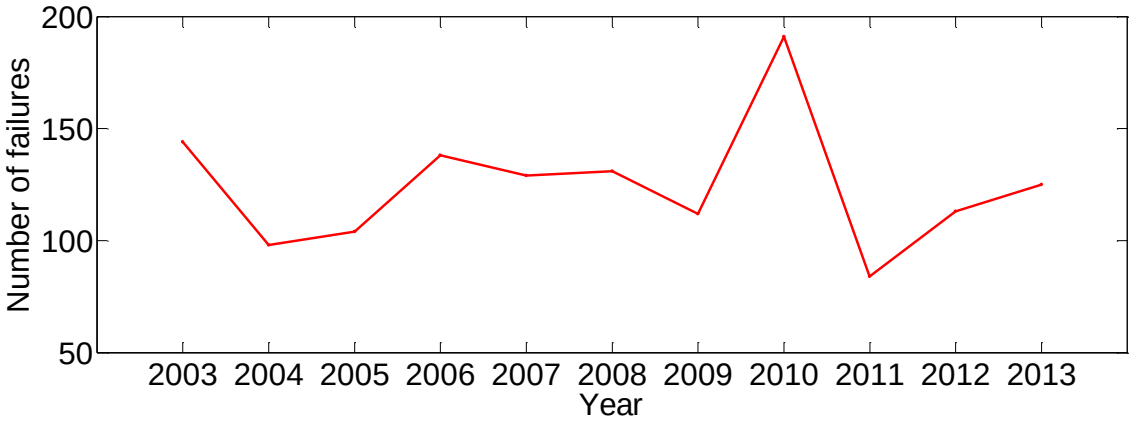

Figure 5.5: Total number of failures per year

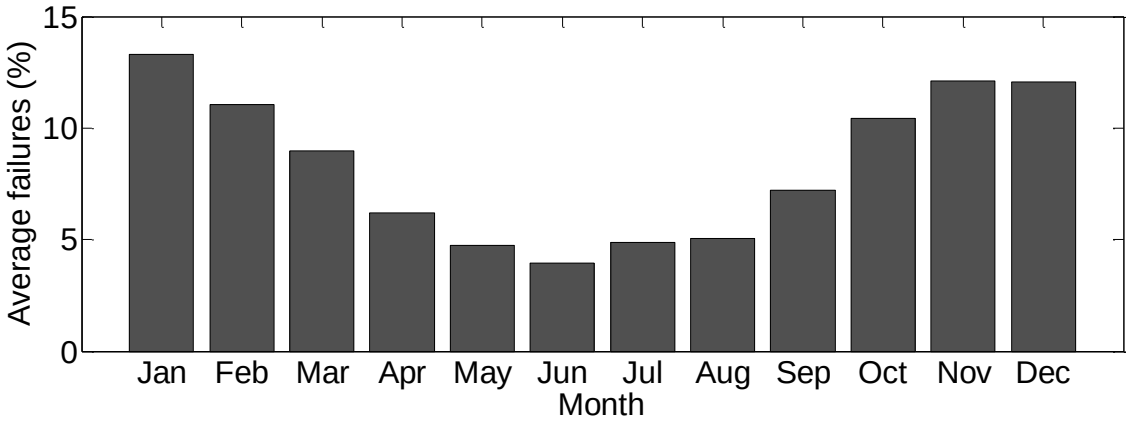

Figure 5.6: Average percentage of failures per month

Figure 5.6 confirms the empirical knowledge that there is an increased number of pipe failures during the coldest months of the year in the UK. The fluctuation of failure frequency versus time (both monthly and annually) justifies the use of weather-related factors which are dynamic for making predictions.

5.2 Data pre-processing

The available data can be broadly divided into: pipe attributes, surrounding environment and broader environment (Figure 5.7). The ‘’pipe-attributes’’ group includes physical and operational factors while the ‘’broader environment’’ group entails environmental factors. The factors falling into the ‘’broader environment’’ group are all dynamic. The ‘’pipe attributes’’ and the ‘’surrounding environment’’ categories entail factors that are either static (i.e. soil type, material, length, diameter) or dynamic (i.e. soil temperature, age, pressure).

Daily climate data were obtained from the British Atmospheric Data Centre and consisted of the minimum air temperature, the maximum air temperature, the average soil temperature and the rainfall observations.

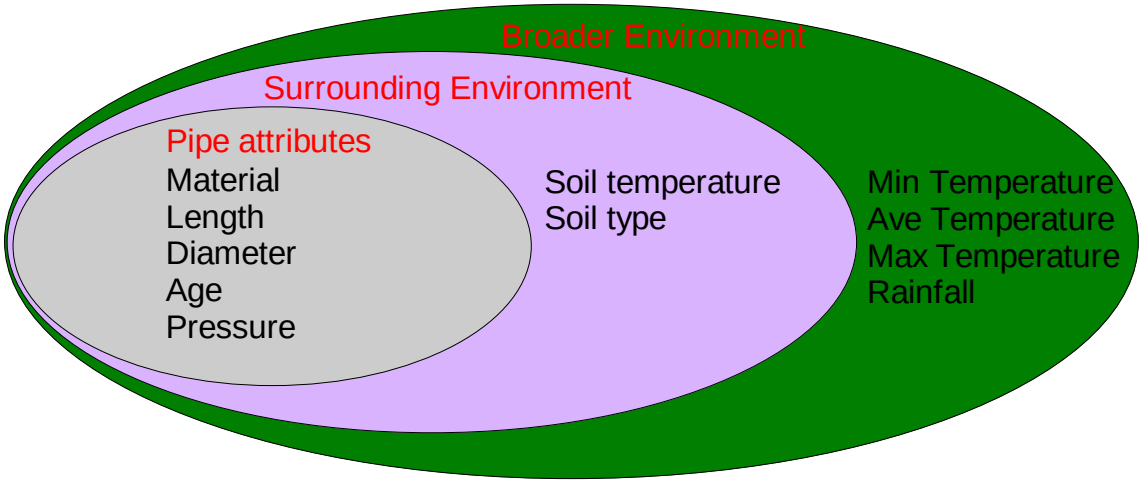

Figure 5.7: Classification of available data

5.3 Preliminary Analysis

The pipe failure is examined on a group level assuming that pipes with the same characteristics are expected to have similar failure rate. The grouping criteria used so far include pipe properties (i.e. material, diameter and age). Tables 5.3 and 5.4 show the average failure rate of the entire examined period versus all the available pipe diameters and installation years. The dispersion of the pipe failure rate in both tables highlights their usefulness in the grouping process.

In addition to the widely-used pipe-intrinsic factors, the use of soil type as *aggregation* criterion is examined. The average failure rate of the CI pipes is 0.203 per km/year but as shown in Table 5.5 it varies significantly depending on the soil type. Examining the failure rate within groups of the same soil type is also a surrogate parameter for data such as soil conditions (e.g. soil moisture) that are difficult to collect and have inherent uncertainty (Phan et al. 2018). Folkman (2018) observed that a CI pipe in highly corrosive soil is expected to have over 20 times the break rate of a CI pipe in low corrosive soil.

Table 5.3 Failure rate vs pipe diameter

Pipe diameter(mm)	**Failure rate**
75	0.241
100	0.216
150	0.157
225	0.080
300	0.049

Table 5.4 Failure rate vs installation year

Installation year	Failure rate
1895	0.213
1905	0.239
1915	0.000
1925	0.112
1935	0.141
1945	0.223.
1955	0.266
1965	0.245
1975	0.216
1985	0.024
1995	0.136

Table 5.5 Failure rate vs soil type

Soil type	Failure rate
C	0.250
O	0.179
T	0.194

The original dataset entailing 18872 individual pipes is converted into a new dataset of 148 homogenous groups. Each of them is made of a unique combination of installation year, diameter and soil type. Note that those 148 homogenous groups do not have the same size (i.e. total length) and number of pipes.

Chapter 6. Results of Predictive Models

6.1 Results for Long-term Predictions

The cluster-based approach was applied for different numbers of clusters (k) and the most appropriate number of clusters was identified by comparing the performance indicators (Figure 6.1) and considering the availability of the data samples. The failure rate which is the targeted output in the analysis has been excluded from the clustering process to avoid circular filling pitfall.

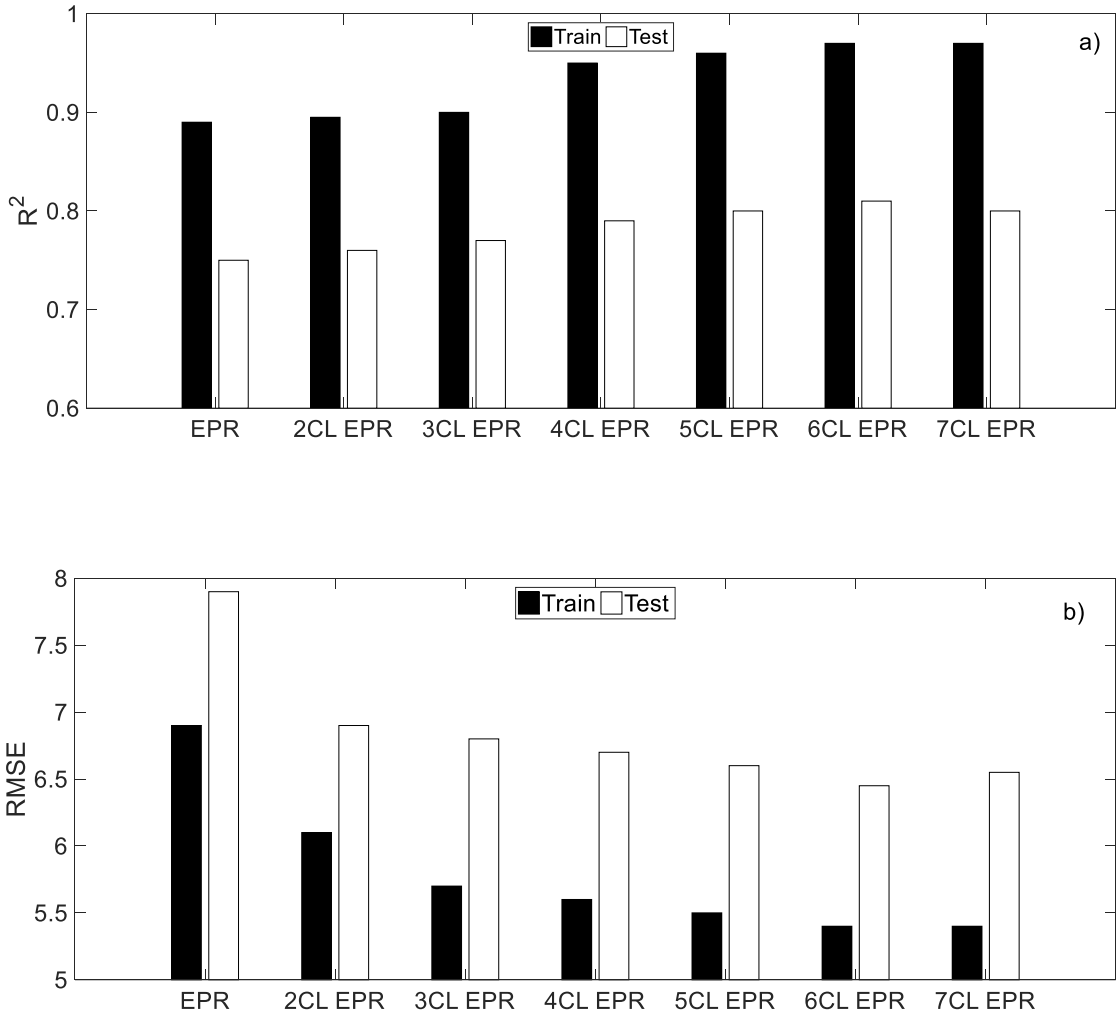

Figure 6.1. Performance indicators of the EPR models (a) R^2 and (b) RMSE

*CL=abbreviation for 'clustered' (e.g. 2CL=two-clustered)

The number of clusters increased until no further improvement was achieved with the test data for both performance indicators and the data within each cluster were too few. The values of the training data in Figure 6.1 are the average values of the 10 iterations of the cross-validation technique. Their comparison shows that the performance indicators are improved by increasing the number of clusters until six clusters. The selection of the number of clusters is case-specific of the examined network. The main purpose is to improve the accuracy of predictions rather than find a general optimal number of clusters. Selecting too many clusters could result in several EPR models that need calibration and reduce the available data to an extent that compromises the predictive accuracy of each model (Osman and Bainbridge 2010).

For comparative purposes, the results obtained from the cluster-based EPR models are compared here with the non-clustered EPR. The results show that both performance indicators for the clustered EPR models are better than the non-clustered EPR approach for all the different number of clusters and for both training and test data. More specifically, the comparison of the six-clustered EPR with the non-clustered EPR shows a significant improvement especially for the test. The RMSE is 6.47 and 7.83 and the R^2 is 0.80 versus 0.75 respectively for the clustered and non-clustered EPR with the test data.

Table 6.1 lists the associated models for the six-clustered EPR and the EPR methods corresponding to one of the ten iterations of the cross-validation. In both approaches, the total group length (L), the diameter (D) and the age (A) of pipes are selected. The two polynomial-terms model also selects the average pressure (P_{ave}) as explanatory variable. The performance indicators of two-polynomial terms six-clustered EPR approach with the test data are 0.82 for the R^2 and 7.64 for the RMSE.

Table 6.1. Obtained formulas for six-clustered EPR and EPR

	Six-clustered EPR		EPR
	One-polynomial-term	**Two-polynomial terms**	
Cluster 1:	$Y=\frac{0.427L^{0.5}A^{0.5}}{D}$	$Y=\frac{0.492L^{0.5}}{D}+\frac{2.141A^{0.5}}{DP_{ave}^{0.5}}$	$Y=\frac{0.015L}{DA}$
Cluster 2:	$Y=\frac{1.196LA^{0.5}}{D}$	$Y=\frac{1.582L^{0.5}A^{0.5}}{D^2}+\frac{0.173A^{0.5}P_{ave}^{0.5}}{D}$	
Cluster 3:	$Y=\frac{0.162L}{D^{0.5}A^{0.5}}$	$Y=\frac{2.043L^{0.5}}{DA}+\frac{0.653P_{ave}}{A}$	
Cluster 4:	$Y=\frac{0.348L^{0.5}A^{0.5}}{D^2}$	$Y=\frac{0.591LA^{0.5}}{D^2}+\frac{1.094A^{0.5}P_{ave}^{0.5}}{D}$	
Cluster 5:	$Y=\frac{2.512LA^{0.5}}{D}$	$Y=\frac{1.741A}{D^2}+\frac{2.159L^{0.5}A^{0.5}}{DP_{ave}}$	
Cluster 6:	$Y=\frac{0.739L}{DA}$	$Y=\frac{2.982L}{D^2A^2}+\frac{0.329P_{ave}^2}{D^2}$	

The returned models indicate a mixed relationship (clusters 1 and 5) between average pressure and failure and a small improvement in the accuracy with the test data in terms of the performance indicators. EPR returns a range of models with varying number of polynomial terms enabling to understand which inputs are physically meaningful and which can be excluded for the sake of parsimony, while simultaneously striving for a degree of generality. For a set of otherwise equivalent models of a given phenomenon one should choose the simplest one to explain a dataset; and also to prevent over-fitting to training data (Young et al. 1996; Crout et al. 2009) which is of concern for medium utilities with limited numbers of recorded failures (Jenkins et al. 2014).

Figures 6.2 and 6.3 indicate a lack of clear relationship between average pressure and failure rate and hence might have limited contribution in explaining the phenomenon. Pressure (both average and maximum) is calculated as the ‘’average value’’ of all the pipes in a group and this mixture might lead to low

impact on pipe failure. Another reason can be that the pressure values are not actual measurements but are the outputs of a calibrated hydraulic EPANET model. Furthermore, the pipe failures were recorded between 2003 and 2013 whereas pressure was taken from EPANET in 2014 and the water utility might have taken measures (e.g. pressure reduction valves) to reduce pressure meanwhile. The average pressure and maximum pressure graphs are very similar. Hence, the maximum pressure is entirely excluded since its selection as explanatory variables would not further improve the accuracy of the models.

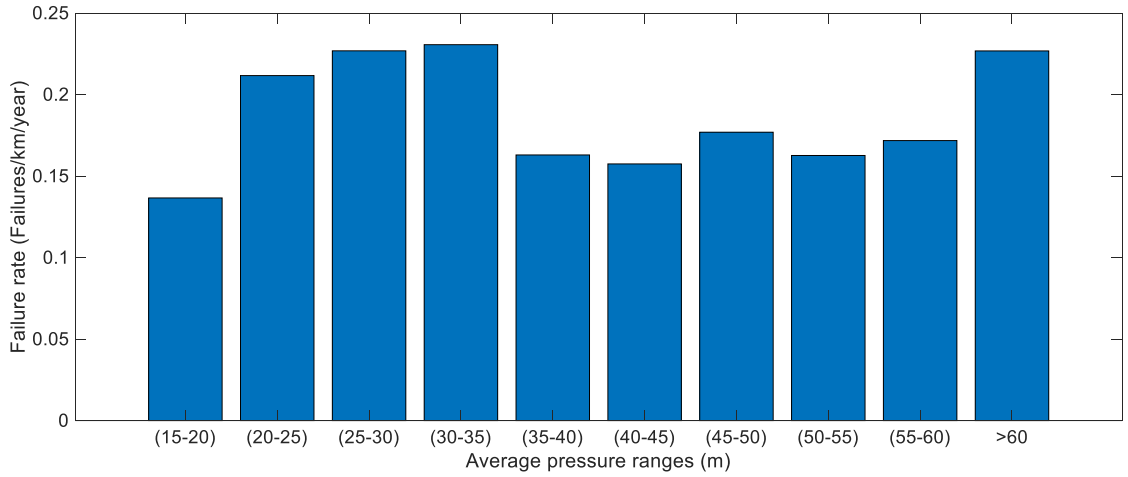

Figure 6.2 Failure rate versus average pressure ranges

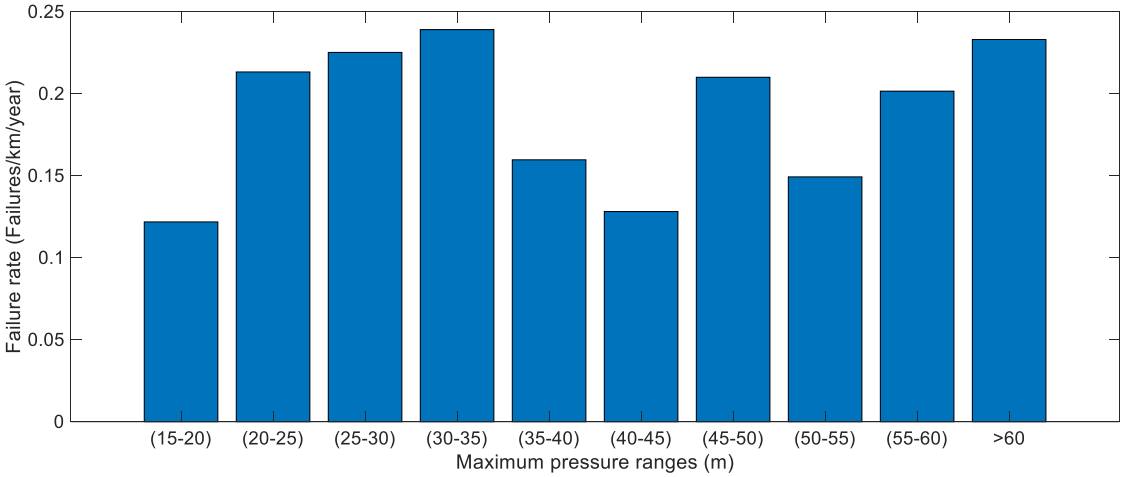

Figure 6.3 Failure rate versus maximum pressure ranges

Shirzad et al. (2014) examined the impact of average pressure on the pipe failure rate of two datasets and found no correlation for one of them. Pressure was not

included in any of the models developed by Park et al. (2001) but the authors attributed it to the fact that the internal pressure data used were for a grid, in which many pipes exist, not for an individual pipe. Andreou (1986) found pressure significant, but of lower importance compared to other physical factors.

Based on the comparison of the accuracy of the single-polynomial term and the two-polynomial terms models and for the sake of parsimony and generality while capturing the physical phenomenon, the one-polynomial model was adopted. It should be noted that this selection does not imply that pressure does not have any impact on pipe failure frequency.

Both single-EPR and clustered-EPR indicate a direct relationship between the total length of the group and the number of failures and an inverse relationship with the diameter which is confirmed in the literature (e.g. Boxall et al. 2007; Berardi et al. 2008; Savic et al. 2009; Xu et al. 2011b). On the contrary, the relationship between failure and age shows some complexity. Four selected models with the six-clustered EPR approach corresponding to clusters 1, 2, 4, and 5 show a direct relationship whereas the remaining two models corresponding to clusters 3 and 6 show an inverse relationship. As shown in Figure 6.4 clusters 3 and 6 entail the oldest pipes. The main reason for this counterintuitive relationship is that the age of many pipes and particularly the oldest ones is much larger than the period their failures were systematically recorded since the examined pipe dataset is left truncated. The left truncation occurs when the pipes were installed before their failures were systematically recorded and the number of failures between installation and the beginning of the monitoring period is unknown (Scheidegger et al. 2013). Several water authorities have a brief recorded failure dataset (e.g. Le Gat and Eisenbeis 2000; Mailhot et

al. 2000; Pelletier et al. 2003; Watson et al. 2004; Vanrenterghem-Raven 2007; Toumbou et al. 2012).

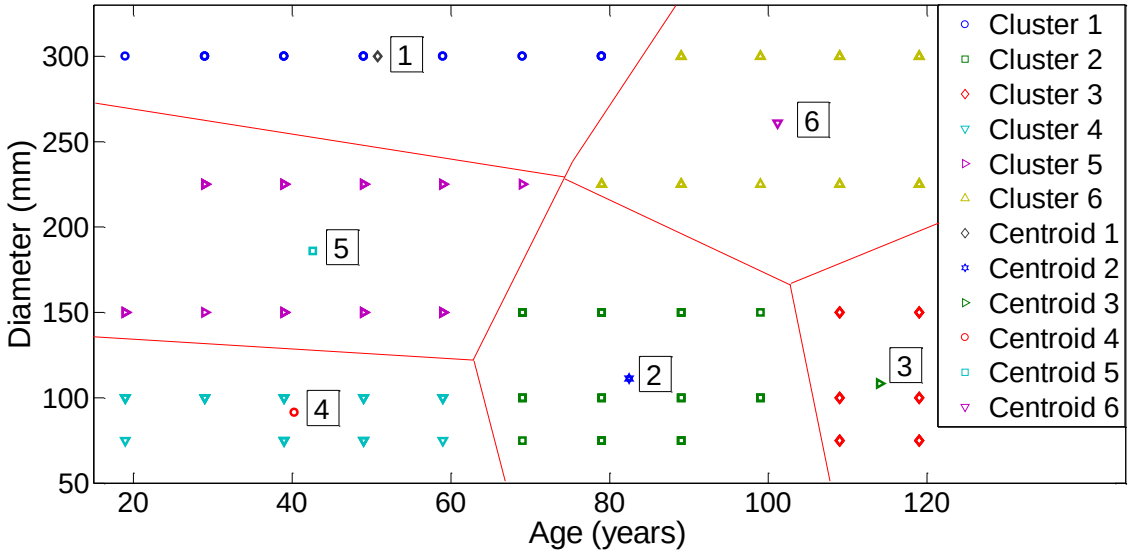

Figure 6.4 Six clusters and the corresponding centroids

Another possible factor can be that only measurable variables are included in the models. Several factors, such as design and construction practice, traffic loads, bedding conditions are not measured and their variation can lead to considerable changes in the subsequent performance of pipes from one age group to another (Boxall et al. 2007). Boxall et al. (2007) has also observed a discrepancy in the association between age and pipe failure. Xu et al. (2011b) examined a brief recorded pipe breakage dataset. They partitioned the pipe database into two clusters of those installed before the beginning of monitoring period and those after the monitoring period. The models obtained show an inverse relationship between pipe failure and age for the pipes installed before the beginning of monitoring period, which constitute the older part of the network.

Comparison between EPR and Six-clustered EPR

Further analysis of this comparison can be seen in Figure 6.5 where the RMSE of the test data is plotted for both methods based on different intervals of the number of pipe failures (the relative frequency indicates the percentage of the 148 homogenous groups with this number of failures). The comparison shows a substantial error reduction for pipe failure events with a large number. In addition, although the improvements of the RMSE for the intervals with a low number of failures is small in absolute terms, the overall model accuracy improvement is significant due to impact on over 50% of the database.

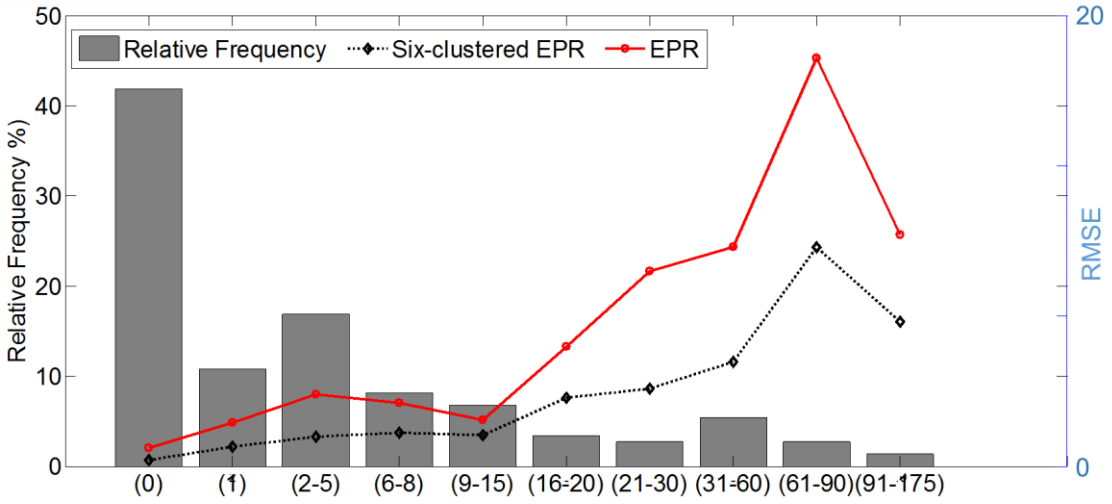

Figure 6.5 Prediction model error for various intervals of number of failures

Spatial variation of pipe failure rate

The *aggregated* EPR models are used to calculate the failure rate (expressed as number of failures/km/year) of individual pipes assuming that within the homogenous groups they share the same failure rate. They are then classified based on their failure rate using the Jenks Natural Breaks method into five ranges as ‘very low’ [0-0.0724], ‘low’ (0.0724-0.1551], ‘medium’ (0.1551-0.2421], ‘high’ (0.2421-0.7663] ‘very high’ (greater than 0.7663]. The user selects the number of ranges and the Jenks method finds the "best" way to split up the ranges. Figures

6.6, 6.7 and 6.8 show this classification of individual pipes with observations, EPR predictions and six-clustered EPR predictions respectively.

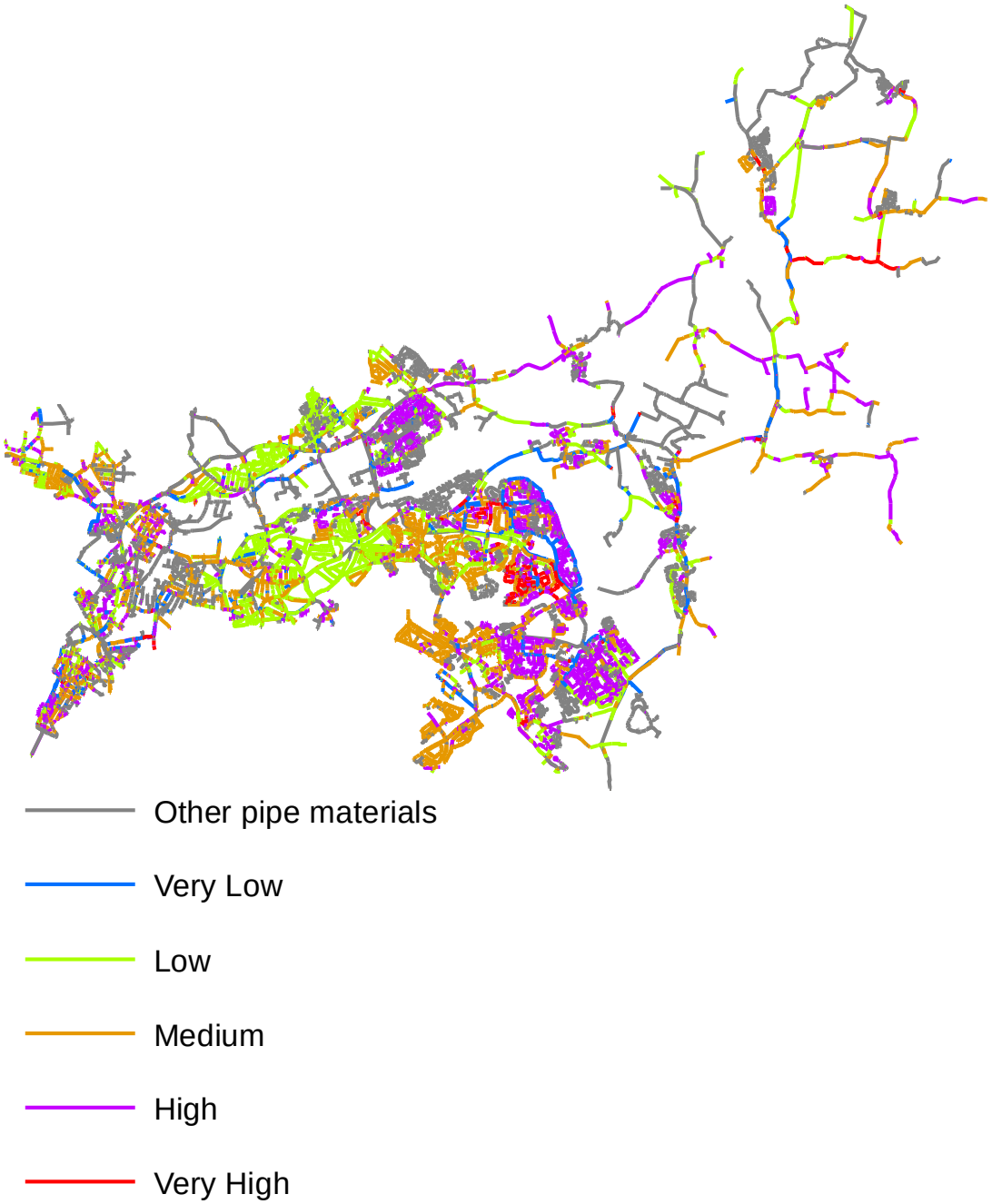

Figure 6.6 Observed pipe failure rate for the entire monitoring period

All the choropleth maps have been created using exclusively test data and they show the average failure rate of the individual pipes for the entire monitoring period (2003-2013).

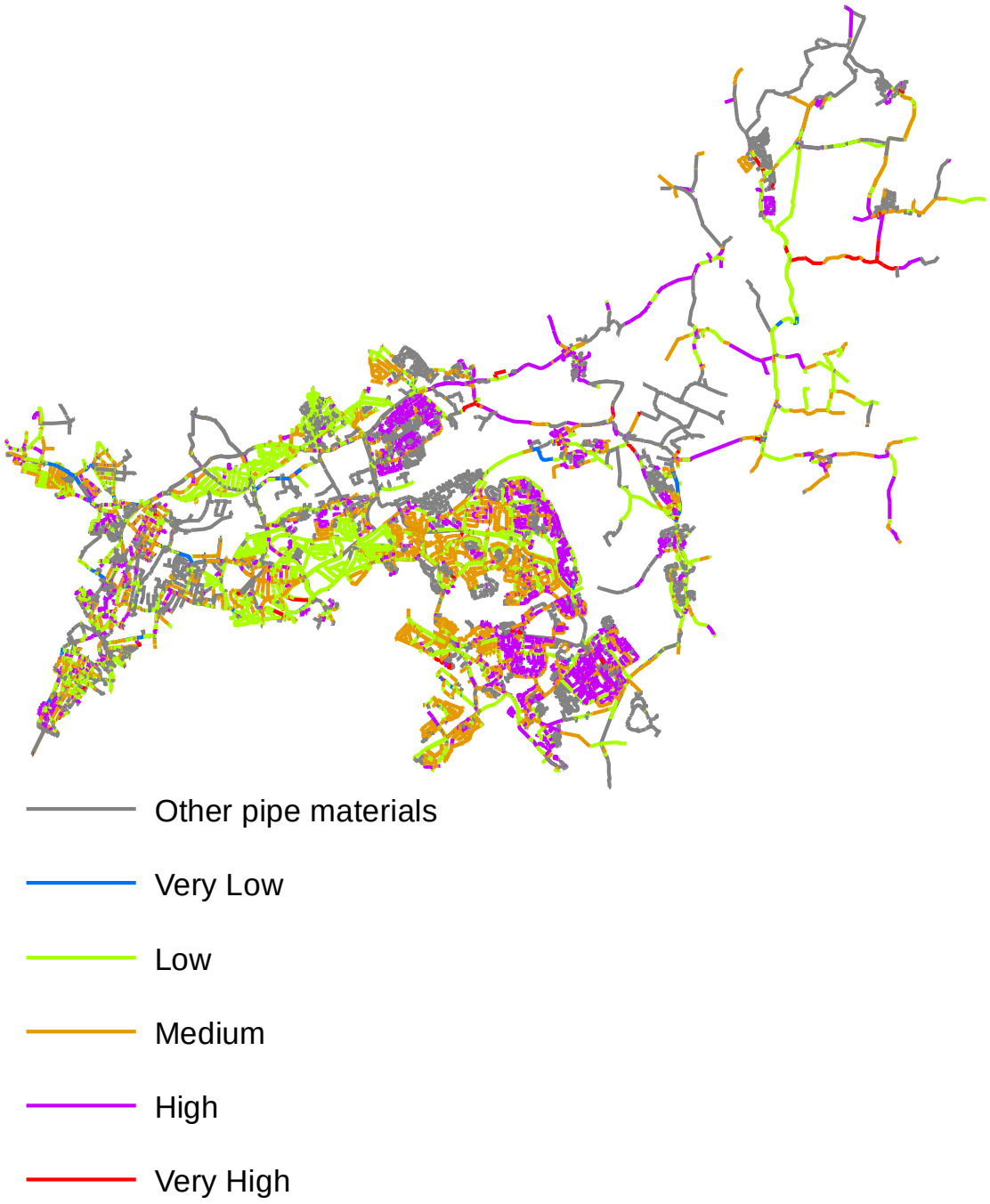

Figure 6.7 EPR predictions of pipe failure rate for the entire monitoring period

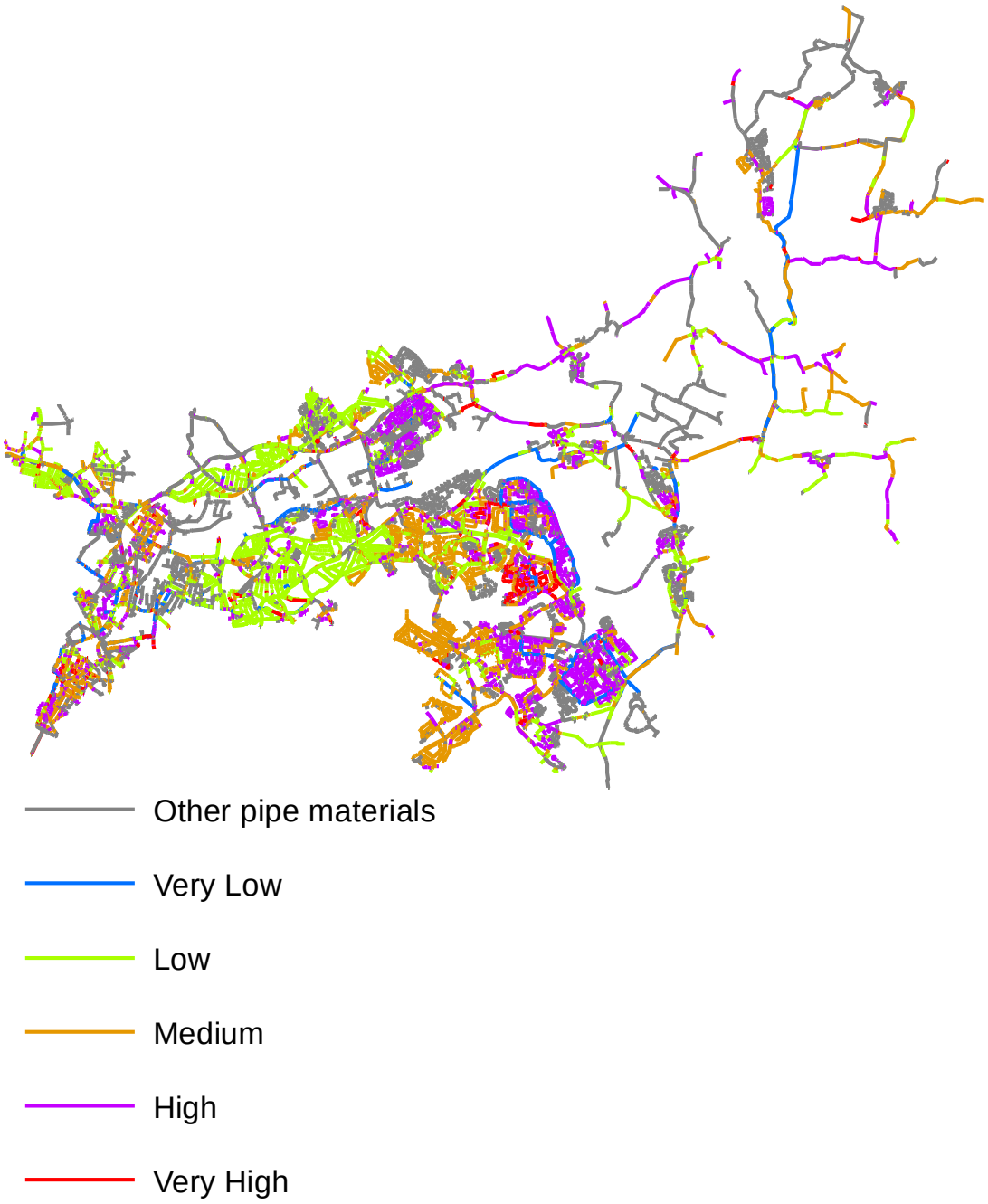

Figure 6.8 Six-clustered EPR predictions of pipe failure rate for the entire monitoring period

Comparison between the accuracy of the two predictive models can be summarised in the pipes that are allocated in the correct range as shown in Figures 6.9 and 6.10. The pipes that are allocated in the correct range are highlighted as green and those that are allocated in a wrong rang are highlighted with red.

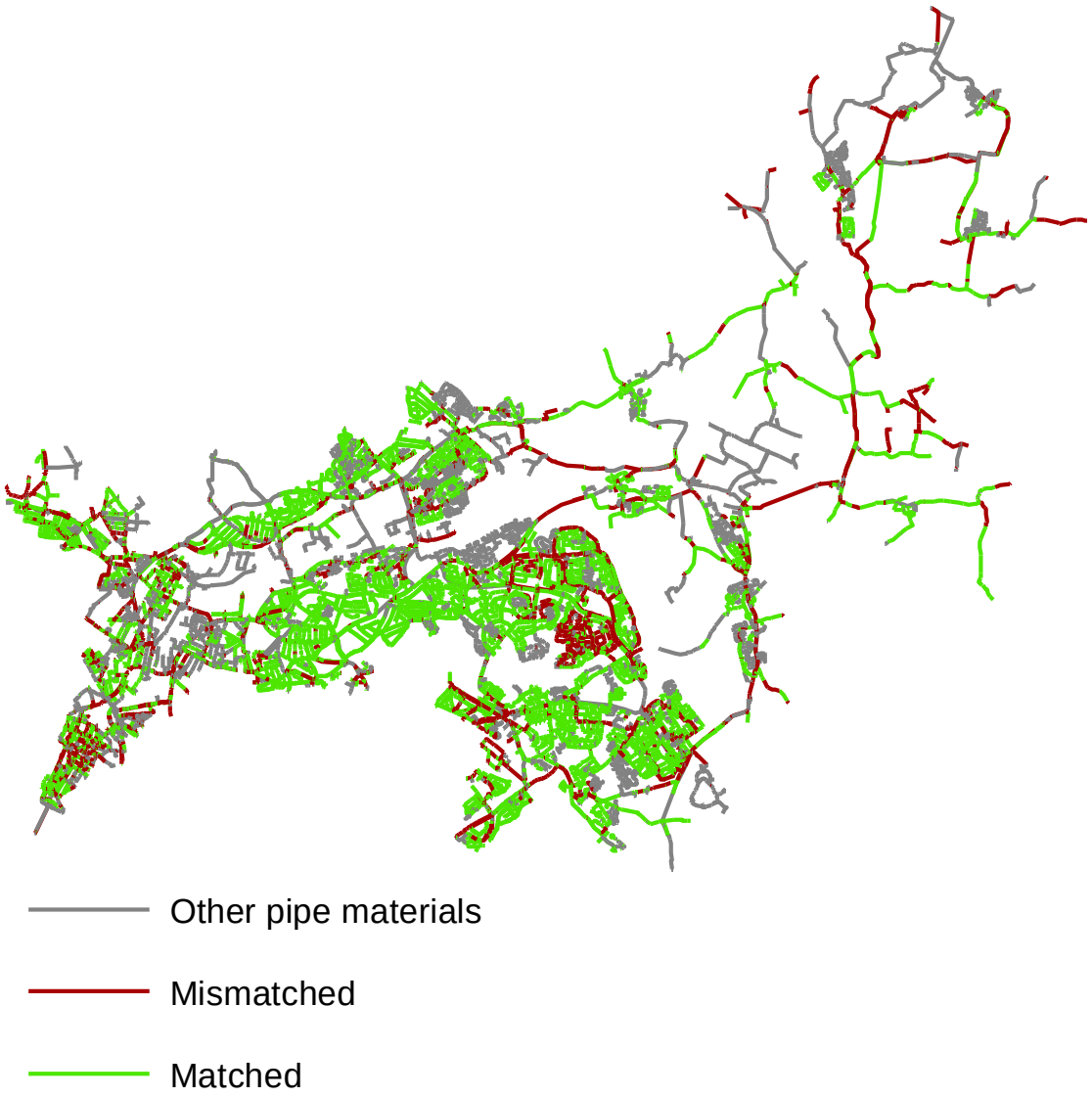

Figure 6.9 Mismatched and matched EPR predictions

Figure 6.9 indicates that there are numerous pipes that are not allocated in the correct range using the single EPR approach. This can be explained by the low accuracy of the non-clustered approach. Figure 6.10 shows that there are less pipes allocated in wrong ranges compared to the six-clusters approach. Direct comparison between the accuracy of the approaches is hard to be done using only Figures 6.9 and 6.10 because the WDN consists of approximately 27,000 pipes (including both the examined CI, the CI that have been excluded and other pipe materials).

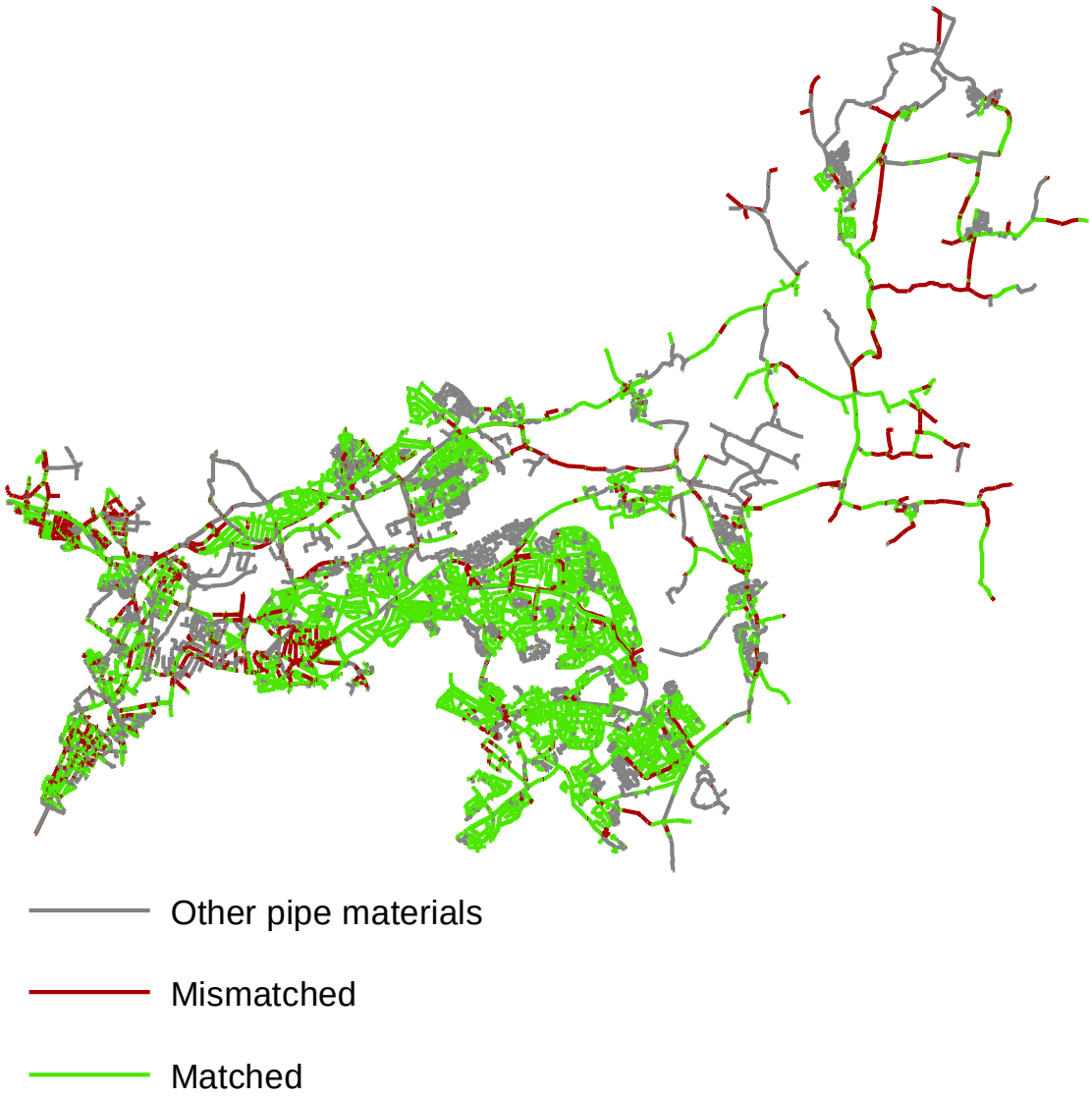

Figure 6.10 Mismatched and matched six-EPR predictions

The improvement in predictions indicated by the comparison between Figures 6.9 and 6.10 is confirmed in Figure 6.11 which shows the pipes that are not correctly allocated with the single EPR approach but are correctly allocated with the six-clustered EPR approach. As shown in Figure 6.11 a significant improvement in pipes allocation is achieved when the examined dataset is divided into six clusters compared to the single EPR method. The clustered-based approach captures the variability of failure patterns better than the single EPR model particularly the very low and the very high failure rates.

Figure 6.11 Improvement in the predictions

6.2 Results for Annual Predictions

Following the proposed the methodology, 11 datasets were created for each of the six clusters corresponding to the duration of the monitoring period (11 years). The methodology resulted in six EPR models each of them corresponding to the training data (2003-2012) of the relevant cluster. Table 6.2 lists the associated models and the R^2 with the test (2013) dataset. EPR selects the 'best set of covariates' meaning a set of covariates that provides close matches between observed and predicted values and at the same time encompasses a minimal number of covariates.

Table 6.2 Obtained models for the six clusters for the annual predictions

Six-clustered EPR	R^2
Cluster 1: C_1=0.001($\mathrm{FI}^{0.5}$)+0.46	0.80
Cluster 2: C_2=2.513($\mathrm{FI}^{0.5}$)+8.52	0.78
Cluster 3: C_3=1.834($\mathrm{FI}^{0.5}$)+5.21	0.84
Cluster 4: C_4=3.425($\mathrm{FI}^{0.5}$)+12.76	0.93
Cluster 5: C_5=1.143($\mathrm{FI}^{0.5}$)+4.05	0.75
Cluster 6: C_6=0.01($\mathrm{FI}^{0.5}$)+1.32	0.86

The selected threshold for the freezing index is 0^0C because it provided the highest correlation in the preliminary analysis. The one-polynomial term EPR model selected the FI whereas minimum air temperature, maximum air temperature, average soil temperature were not selected. The mathematical relationship indicates that lower temperatures and consequently higher values of FI increase the number of failures.

The highest correlation (the higher values of R^2) is observed for cluster 4 which includes the small diameter pipes (Figure 6.3). Fuchs-Hanusch et al. (2011) have observed a dependency between failure frequency and FI especially for pipes with up to 200 mm diameters. Bruaset and Sægrov (2018) also observed a higher correlation between failure and temperature during winter months for the smaller CI pipes.

As shown in Figure 6.12 there is a clear association of FI (solid black line) variability and the pipe failure (bars) whereas this correlation can't be observed for any of the values of the temperature (minimum temperature is the red line, maximum temperature is the purple line, average temperature is the blue line, soil temperature is green line). The temperature-related candidate explanatory

variables are calculated as average values, and this might act dissuasively in capturing periods with severe conditions. For example, winter/early spring 2018 have been very cold while summer has been very hot; hence the average temperature for 2018 may fail to highlight the severity of the cold period.

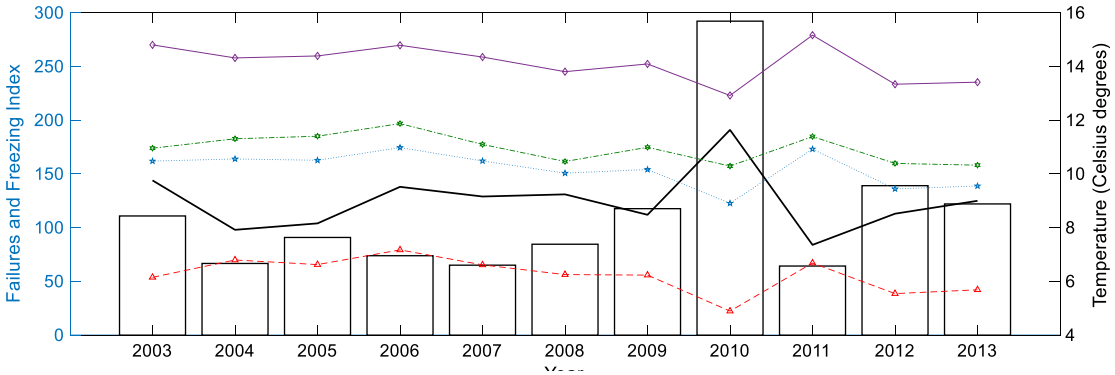

Figure 6.12 Number of failures and weather conditions per year

Figure 6.13 shows the predictions vs the observations for all the clusters for test dataset, year 2013. The EPR models exhibit a high accuracy in predicting the number of failures for clusters 1 and 6. The difference between the predictions and the observations is low for rest of the clusters. The absolute difference between observations and predictions for clusters 1 and 6 tends to zero whereas it varies between 3 and 6.5 for the rest of the clusters.

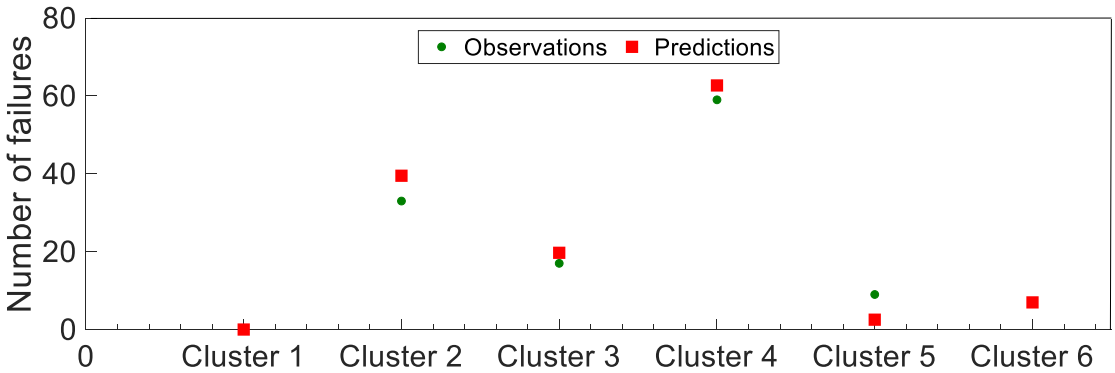

Figure 6.13 Predictions vs observations for 2013

Spatial variation

The individual pipes were classified using the Jenks Natural Breaks method into five ranges as ‘very low’ [0-0.091], ‘low’ (0.091-0.236], ‘medium’ (0.236-0.472], ‘high’ (0.472-0.75] and ‘very high’ [greater than 0.751] as shown in Figures 6.14 and 6.15 (observations and predictions respectively).

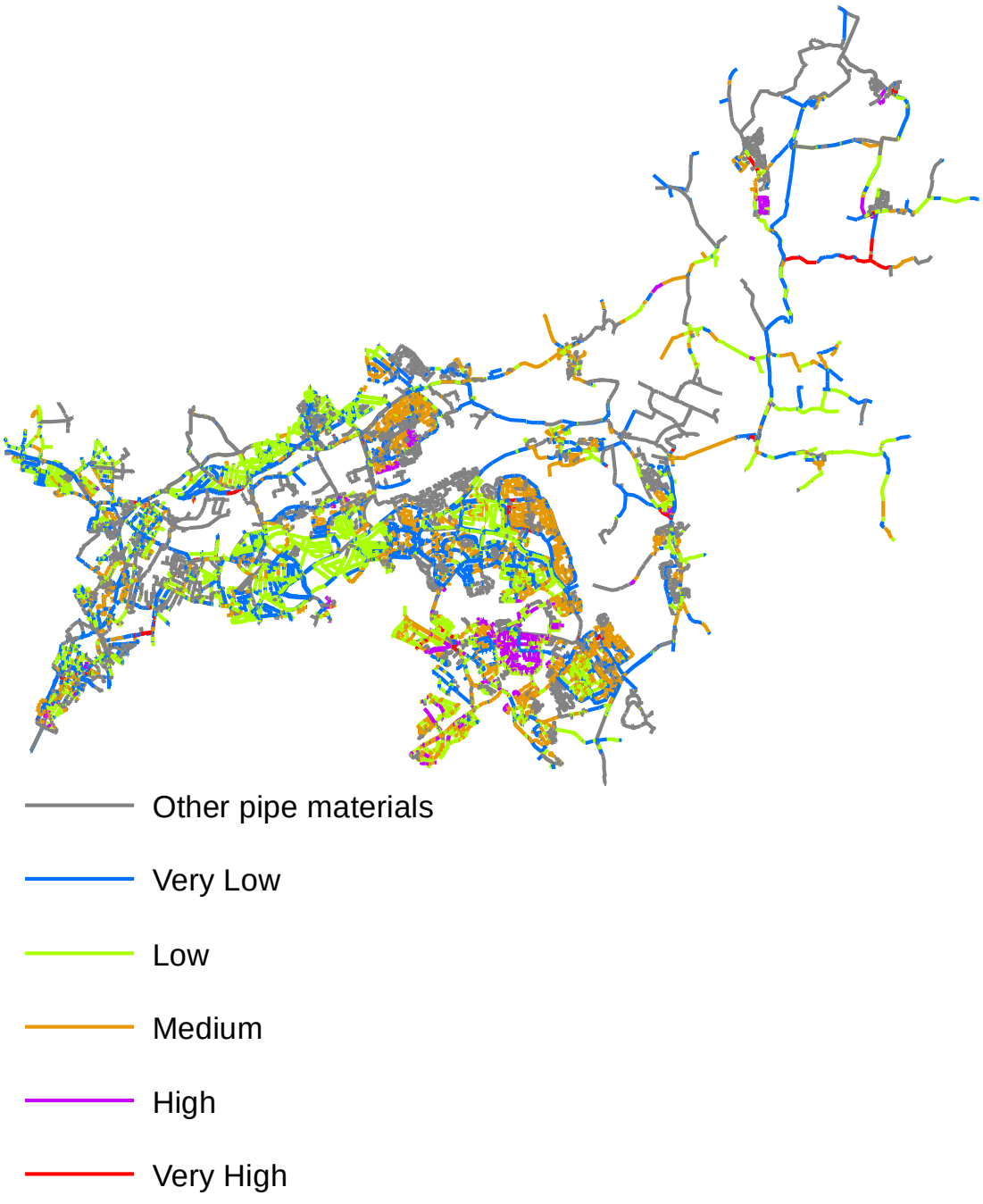

Figure 6.14 Observed pipe failure rate in 2013

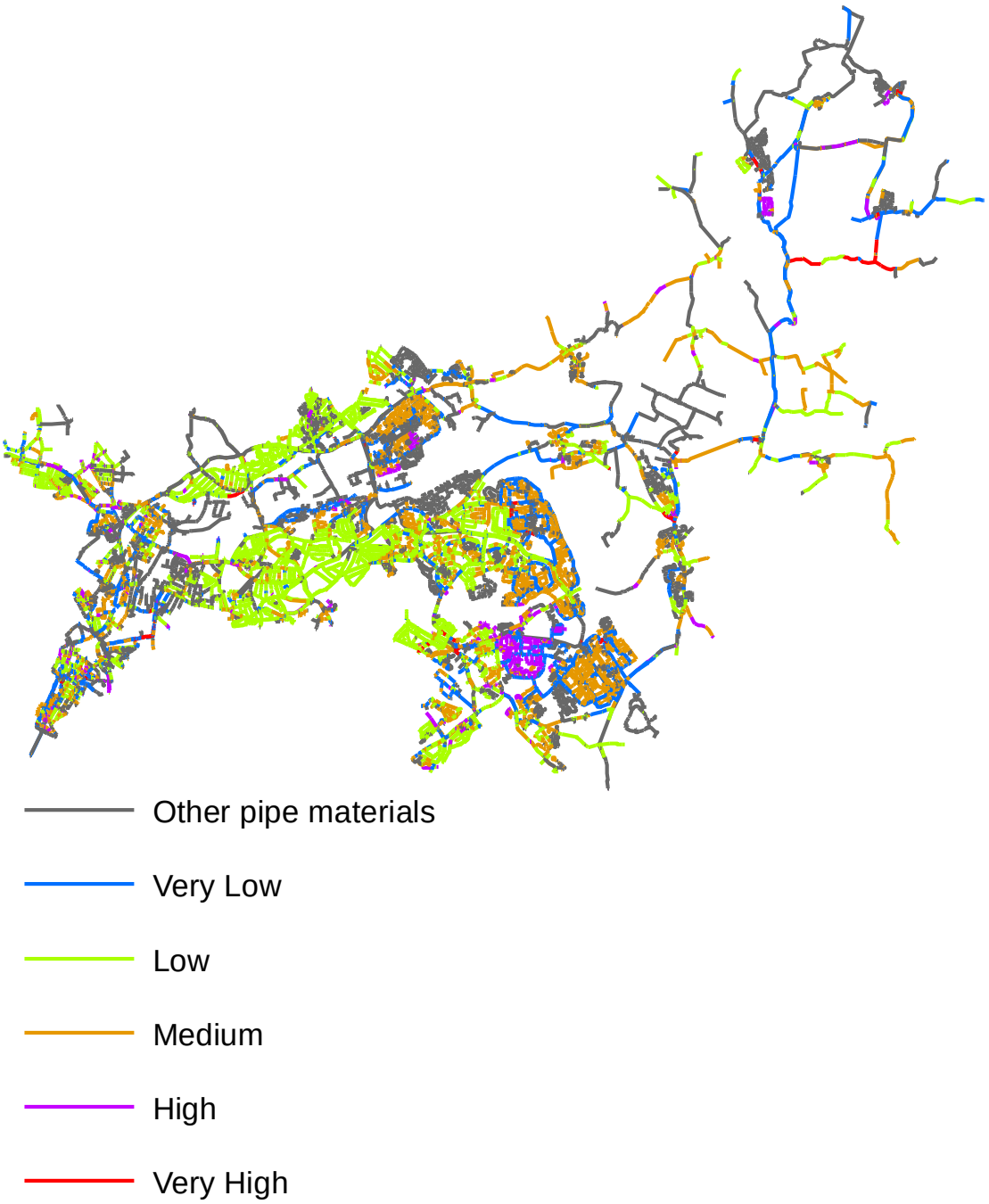

Figure 6.15 Predicted pipe failure rate in 2013

The accuracy obtained with the predictions is 46%, 73%, 78%, 87% and 76.34% for the ‘very low’, ‘low’, ‘medium’, ‘high’ and ‘very high’ failure rates respectively. The predictions have a high accuracy for most of the failure ranges (‘low’, ‘medium’, ’high’ and ‘very high’). The lowest accuracy is achieved for the pipes with a ‘’very low’’ failure rate and is attributed to the fact that some homogenous

groups have experienced zero number of failures. The predicted failures for each cluster are distributed to the homogenous groups proportionally to their length value leading to a slight overestimation. Figure 6.16 shows the predictions after including the physical factors.

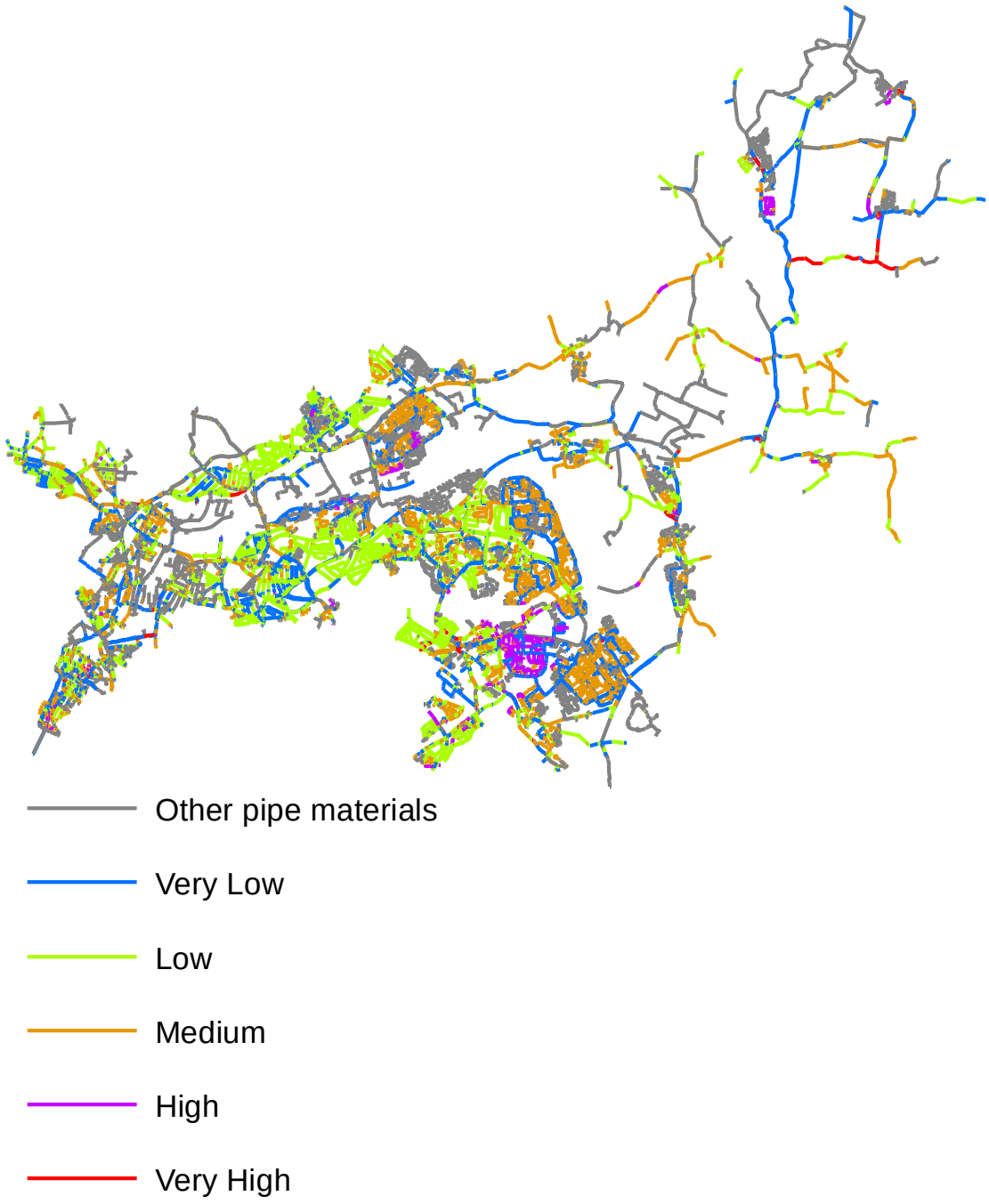

Figure 6.16 Predicted pipe failure rate in 2013 including physical variables

Figure 6.17 compares the accuracy of the predictions when merely environmental variables are used and when they are combined with the physical variables. The inclusion of the physical factors increased the accuracy of the predictions for the majority of the ranges. The highest improvement is observed for the 'very low' range for which shifted to 69%.

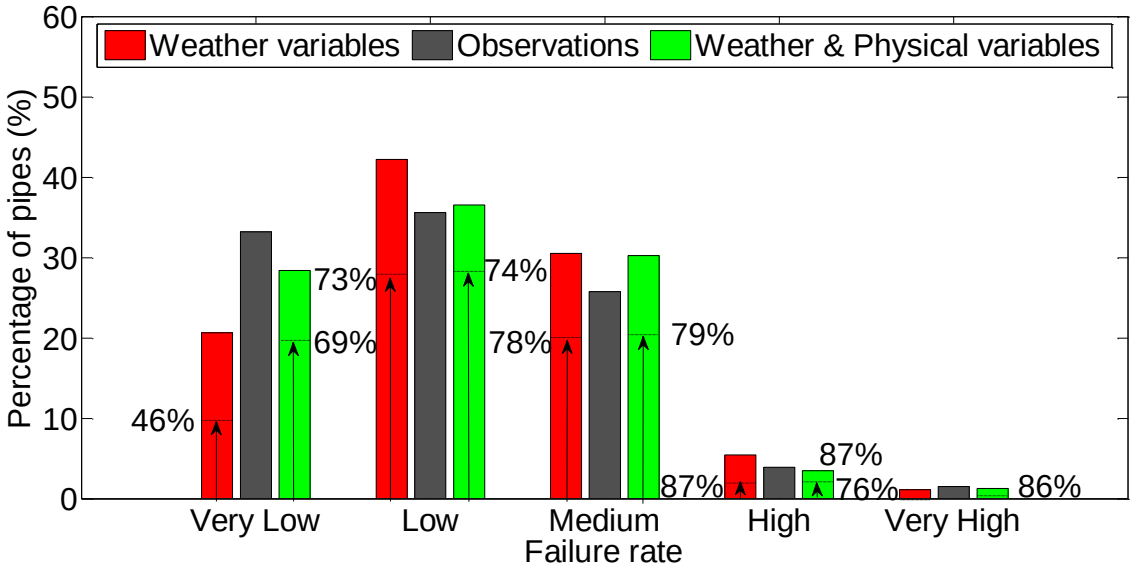

Figure 6.17 Percentage of pipe failure rates for predictions and observations in different ranges; note the percentage next to the bars is the percentage of the correct predictions for each range

The main drawback of the approach is that it requires next year's weather conditions (which need to be forecasted) to make predictions introducing a degree of uncertainty. To overcome this uncertainty, the use of lagged explanatory variables, i.e. use of previous and this year's weather data to make predictions for the next year was attempted. This attempt led to low accuracy since as indicated in Figure 6.12 the weather data are in line with that year's failures. A representative example is year 2010 when the peak of failure frequency coincides with the greatest value of FI.

6.3 Results for Short-term Predictions

This approach is implemented for the entire network and not separately for each cluster due to the small number of failures in some of them. The temperature variation can occur relatively quickly whereas the potential pipe failure because of that might take longer and hence the duration of the inputs/outputs are fixed in this analysis. Exhaustive trials (different combinations of inputs and outputs) were conducted. The original dataset is split into 10 equal size folders using the cross-validation method. Those folders are used for training (i.e. 7 folders), validation (i.e. 2 folders) and test (i.e.1 folder) purposes respectively.

The model's responses are compared to a set of threshold values and the generated pairs of TPR/FPR are used to plot the ROC curve (Figure 6.18) for four combinations of inputs/outputs that provided the highest accuracy. Due to the large number of trials (all the combinations from one up to seven days as input and from one up to seven days as output), it is not feasible to plot all the results. The selected model (red curve) considers four days as input and the following two days as output. The first input is the set of variables for the first four days and the output is the occurrence of failure(s) in the fifth and sixth day. Respectively, the second input is the set of variables from the second up to the fifth day, while the output is the failure in the sixth and seventh day. The black curve model considers three days as input and two days as output, the green curve model considers four days as input and one as output while the blue curve considers three days as input and one day as output. There is also anecdotal evidence from another UK water company that the highest correlation is observed between four days as input and one day as output.

The selected threshold with the lowest Euclidean distance from the optimal point is 0.538 obtained for the model with the four days/two days as combination of

inputs/output. As shown in Figure 6.18 the majority of the non-failures are correctly identified (the FPR is 0.87) as such while a similar conclusion can be derived for the failures despite the lower accuracy (the TPR is 0.72). The value of AUC which is used as a measurement of model's performance is 0.814 indicating that the model has a very good accuracy.

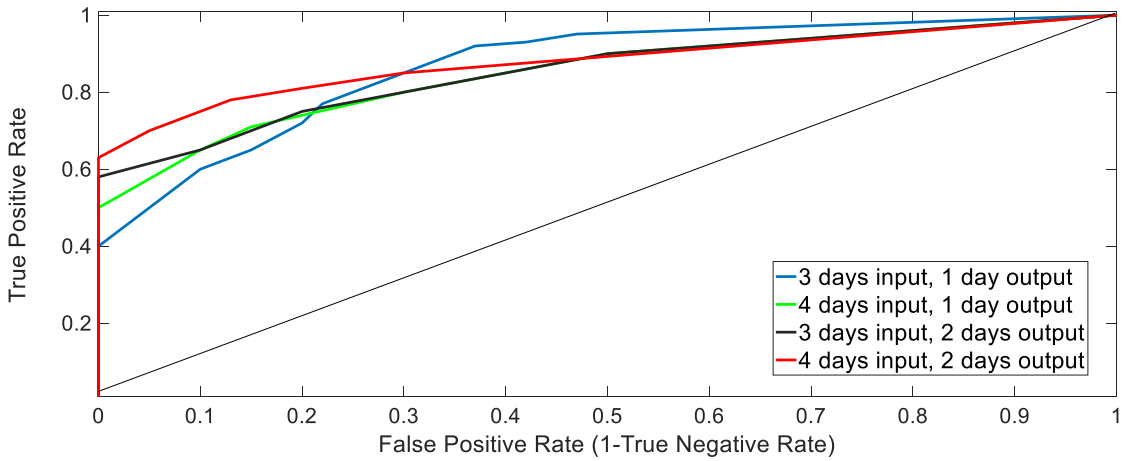

Figure 6.18 ROC of the binary ANN model

As indicated by the outputs of the preliminary analysis (Figure 5.6) most of the pipes burst happen during the coldest months. Water pipes burst because the water inside them expands as it gets close to freezing, and this causes an increase in pressure inside the pipe. When the pressure gets too high for the pipe to contain, it ruptures. This phenomenon is elucidated by the fact that when water cools the molecules slow down. This slowing down allows the molecules to get closer together and increases the density of the liquid. The cold spell normally lasts for a couple of days and there is an increased number of failures during this period or straight after it end since there might be a time lag between temperature drop and pipe failure. Another reason explaining the pipe failure can be the freeze and thaw in soil which causes soil movement. This in turn results in lack of continuous support beneath the pipes creating bedding stresses.

Due to the black-box nature of the ANN model, the relative significance of each input must be estimated. Table 6.3 is a column matrix with the weight of all the inputs which is a surrogate of their influence on the model's response The negative connection weights of the temperature related inputs indicate an inverse relationship (lower temperatures cause more failures).

Table 6.3 The weight of the inputs on the response

Input	Weight
FI_{t-2}	15.75
FI_{t-1}	13.03
FI_{t-3}	12.35
$MinT_{t-2}$	-11.41
$MinT_{t-4}$	-11.02
Max_{t-2}	-10.61
$Soil_{t-2}$	-10.26
$MinT_{t-1}$	-9.55
$Soil_{t-4}$	-9.54
$MaxT_{t-4}$	-9.49
$Soil_{t-3}$	-8.69
$MaxT_{t-1}$	-8.35
$Soil_{t-1}$	-7.66
$MaxT_{t-3}$	-7.53
$MinT_{t-3}$	-7.33
FI_{t-4}	6.32

The FI is shown to be the most influential factor which is linked to the fact that most of the failures occur in the coldest months (Figure 5.6) when pipes are subject to frost actions which significant increases the loads on them (Morris 1967; Monie and Clark 1974; Smith 1976; Habibian 1994; Rajani et al. 1996). The frost load is influenced by frost penetration, trench width, soil type, soil stiffness,

frost heave of trench fill and side fill as well as the interaction at the trench backfill–side fill interface (Rajani and Zhan 1996). The increased number of failures has been correlated with colder winter months (Vreeburg et al. 2013; Kutyłowska and Hotloś 2014) while O'Day (1987) found that 40% of the breaks occur during the three winter months.

The developed model was successfully applied to a holdout sample, demonstrating that the ANN 'learned' the breakage patterns rather than memorised them. The proposed approach uses existing data (that do not have to be predicted) as input which reduces significantly the uncertainty of the output. The predictions can be made based on recorded weather conditions enabling the water utility to adjust their immediate operational planning strategies to accommodate possible increases in pipe failure occurrence.

Chapter 7. Results of Impacts Assessment

7.1 Introduction

Performance measures are indicators that describe the behaviour of a system in terms of its tangible operational characteristics (Mansoor 2007). Assessment of the performance of a WDN is a complex process because many issues should be considered (e.g. variations in demands, reliability of individual components and their locations, fire flow requirements and their locations) (Kalungi and Tanyimboh 2003). Further complications stem from the fact that it is difficult to define useful performance measures and establish what acceptable levels for these parameters are (Kalungi and Tanyimboh 2003). For a WDN, performance indicators quantify its behaviour mainly based on the nodal outflows, supply pressure at consumer outlets, supply interruptions, amount of leakage and water quality issues (Mansoor 2007). The volume of the water loss depends on the characteristics of the network (e.g. the pressure and the flow rate in the network) and other factors, such as the company's operational practice and the level of technology and expertise applied to controlling it (Farley and Trow 2003; Rahmani et al. 2015). The severity and duration of disruption as a result of pipe failure depend on the network layout and structure of cycles and loops as alternative supply paths (Yazdani and Jeffrey 2011; Singh and Oh 2015; Di Nardo et al. 2017a).

The distribution pipes (which is the case in this thesis) usually range in size between 50 to 300 mm in diameter and are typically laid within the road and have failure costs, due to the vigorous and challenging repair techniques and the number of customers impacted (Ward et al. 2017).

7.2 Performance Indicators

The proposed methodology for impacts assessment was implemented for all the examined CI pipes (i.e. 18872) twice assuming that the pipe occurs either at 4am or 7am. In both cases the simulation period is 24hrs commencing from the failure occurrence. The diurnal demand variation in nodes, the water level in storage tanks, and the valve/pump control settings are considered over this predefined simulation period by changing the parameters of the connected FCVs and emitters according to the current values of the desired demands in deficient nodes.

The majority of the examined CI pipes (i.e. 18872), in both scenarios, resulted in low impacts (i.e. low ratio of unsupplied demand and small number of affected nodes). This can be attributed to the fact the examined network is a large real-life network and a single failure of a distribution pipe is likely to cause pressure deficient conditions in a small part of it. Marlow et al. (2015) have also observed that for small-diameter CI pipes, the consequences of failure are often relatively low since the surrounding nodes are mainly affected, whereas performance elsewhere in the WDN is mostly satisfactory (Tanyimboh et al. 2001). In each system, only a few pipes represent the most critical hydraulic links, e.g. trunk mains service connecting reservoirs to the rest of the system (Diao et al. 2016). In a branched (no loop) network with a single source of supply without any other service reservoir locations, the pipe failure will disconnect all consumers downstream resulting in a total and immediate loss of service (Germanopoulos et al. 1986; Jowitt and Xu 1993). In other circumstances, failure may manifest itself in a less complete way; (i.e. consumers may be faced with a drop of pressure to levels below which full demands cannot be met) (Jowitt and Xu 1993). The complex structure of WDNs allows them to recover from failures, exploiting

the topological redundancy provided by closed loops, so that the flow could reach a given demand node through alternative physical paths (Ormsbee and Kessler 1990; Goulter et al. 1999; Gupta et al. 2014). This redundant design enables the system to overcome local pipe failures, and, together with pipe diameters larger than those strictly necessary to fulfil the design pressure at the nodes (Todini 2000; Babayan et al. 2007; Di Nardo et al. 2017b; Giudicianni et al. 2018; Zarghami et al. 2018). The ability of a network to respond to the failure of one of its pipes does not depend only on redundancy conditions in the immediate vicinity of the failure (i.e. at the nodes at either end of the failed link) (Awumah et al. 1990). Alternate paths for supplying nodes may originate some distance from the nodes in the immediate vicinity of the link failure (Awumah et al. 1990). These alternate paths may also not use any links in the vicinity of the failed link (Awumah et al. 1990; Yazdani and Jeffrey 2011). The reservoir elevations, the distance from the source node and the elevation of the nodes are also critical factors that enable the network to work under spatial and temporal pressure deficient conditions (Sivakumar and Prasad 2014).

The loss of performance may be progressive. A pipe failure will initially lead to uncontrolled flow from the network and in general drop in pressure. Service will be totally lost if any shortfall between inflow (supply) and total outflow (uncontrolled flow plus nodal consumptions) exhausts reserved supplies (e.g. service reservoir storage) (Germanopoulos et al. 1986). Isolation of the failed element will allow the loss of water to be stemmed, network heads to be partially restored and the network to be operated at a more stable (but maybe unacceptable) state (Jowitt and Xu 1993). Eventually, the pipe repair will allow a return to normal operation with adequate heads and service reservoir levels within normal operating bounds (Germanopoulos et al. 1986).

This thesis examines a single pipe failure since most of the days there is only one (73.25% of the days there are not any failures, 21% of the days only one failure occurred and more than one occurred on 5.75% of the days). The probability of simultaneous failure of pipes becomes exceedingly small as the number of the failing pipe increases and even the probability of the simultaneous failure of two pipes is quite small for most systems (Agrawal et al. 2007; Gupta et al. 2014). The combinations of possible multiple failure scenarios grow exponentially as the network becomes larger (Berardi et al. 2014) in parallel with possible inconsistencies and uncertainties associated with hydraulic simulations (Gupta and Bhave 1994; Herrera et al. 2016).

Due to the huge number of scenarios (i.e. 18872x2) it is not feasible to present all the results. Therefore, the five pipe failures (i.e. 5x2 scenarios) that resulted in the highest ratio of unsupplied demand have been selected in the following figures to illustrate the impacts versus time. The diameter of these five pipes is 300mm, which is the largest in the examined dataset. Larger diameter pipes tend to have more severe impacts (greater water loss) since they convey more water. Figures 7.1 & 7.2 show the ratio of unsupplied demand for the 24hrs simulation period, which ranges between 4% and 9% for most examined scenarios. Both scenarios show a relative constant ratio of unsupplied demand for the 24hours period. In case of one pipe failure (scenario 4 at 4pm) the ratio of unsupplied demand exceeds the average value and can be linked to the proximity to two pumps. Normally, at night-time, when demand for electricity is low, the tariff is also low (Bunn and Reynolds 2009). Water utilities are in an advantageous position to maximize pumping to storage during off-peak hours (i.e. night-times) when energy rates are lowest. (Brion and Mays 1991; Bunn and Reynolds 2009; Bonvin et al. 2015).

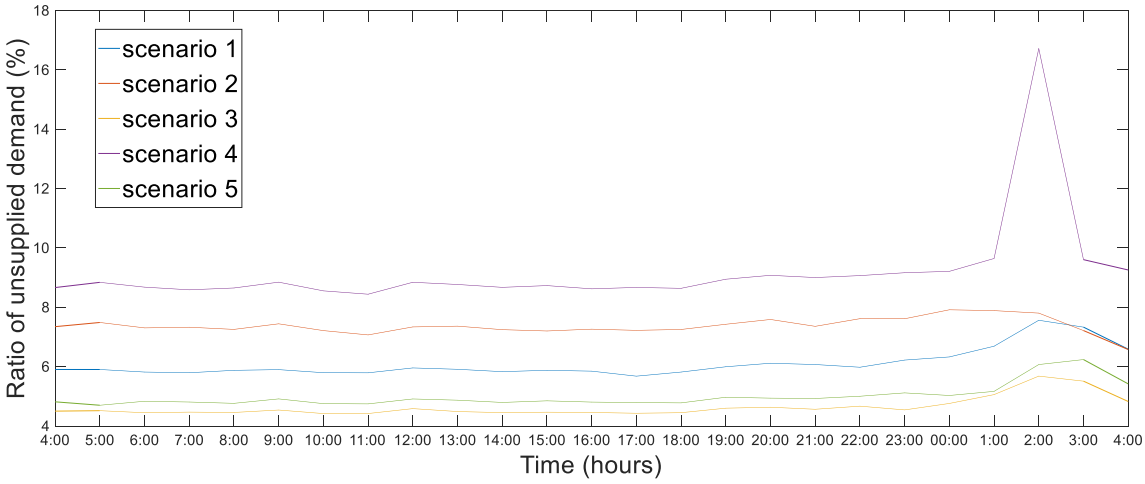

Figure 7.1 Ratio of unsupplied demand for 24hrs after pipe failure at 4am

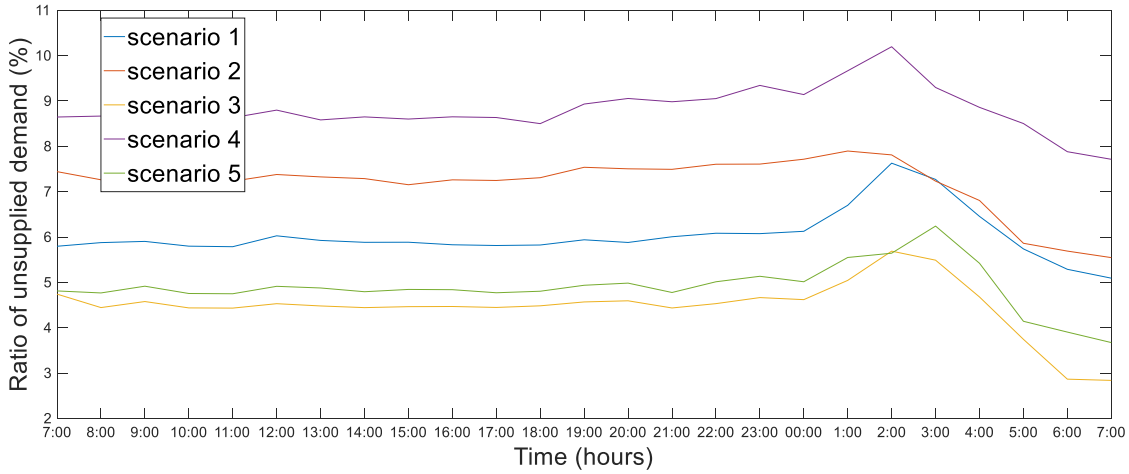

Figure 7.2 Ratio of unsupplied demand for 24hrs after pipe failure at 7am

A pipe failure originally leads to uncontrolled flow from the network and in general drop in pressure in a number of nodes. Service might be totally lost if the inflow is less than the total outflow (uncontrolled flow plus nodal consumptions) exhausting the reserve supplied (e.g. service reservoir storage). In the examined network there are 8 reservoirs (and 2 tanks) and therefore, the increase in the unsupplied demand is relatively small.

The number of nodes experiencing entire loss of supply are shown in Figures 7.3 and 7.4 while those with partly supply are shown in Figures 7.5 and 7.6 further

interpreting the above results. It is shown that there is a similar trend (e.g. (scenario 4 at 4pm) for both the nodes with zero and partly supply. The fluctuation in the number of nodes affected is similar with the ratio of unsupplied demand and remains mostly constant. Despite the similar trend, there is a significant difference between the absolute numbers of nodes affected totally and partially, since in most scenarios there is a small drop of pressure (actual pressure is less than the required but still above zero) leading to partial supplied demand. Hence, the number of nodes with zero supply is very low.

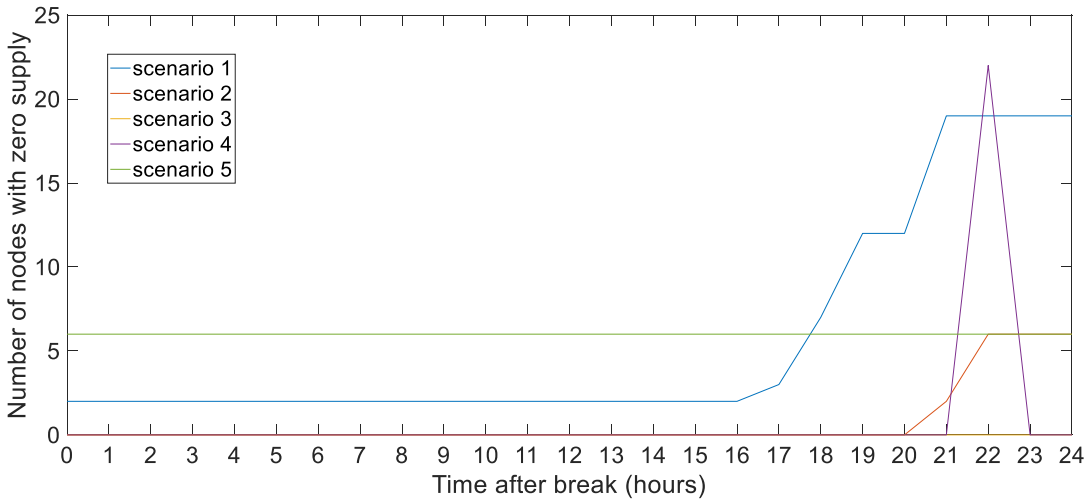

Figure 7.3 Number of nodes with zero supply for 24hrs after pipe failure at 4am

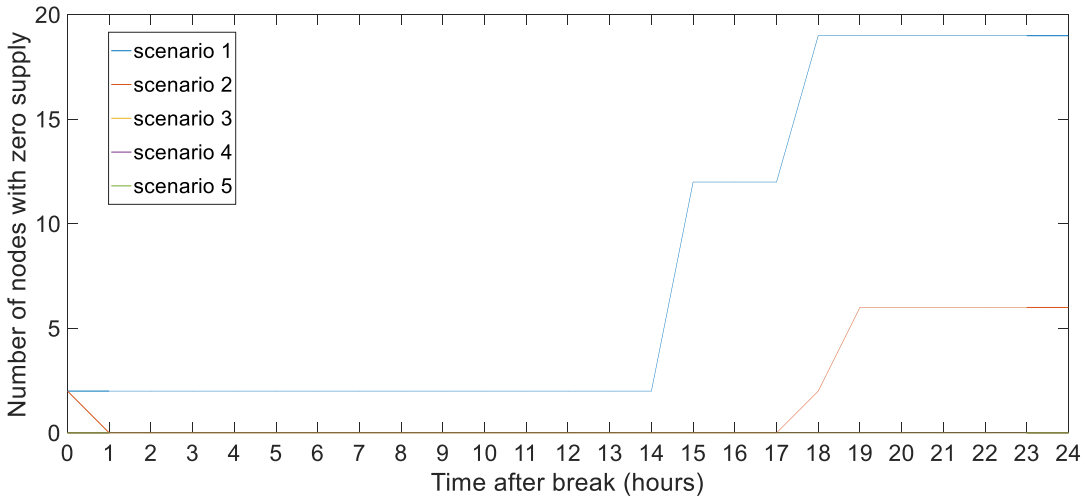

Figure 7.4 Number of nodes with zero supply for 24hrs after pipe failure at 7am

Figures 7.3 and 7.4 show an ascending trend for scenario 1 late in the evening which is associated with the peak in the demand (Figure 4.2). Despite the increase compared to the number of nodes with zero supply, those with partial supply are still very few (maximum 0.009% of the nodes) considering that the network entails approximately 33,000 nodes. Overall the small number of nodes (both zero and partial supply) experiencing pressure deficient conditions indicates that the surrounding nodes are mainly affected, whereas performance elsewhere in the WDN is mostly satisfactory.

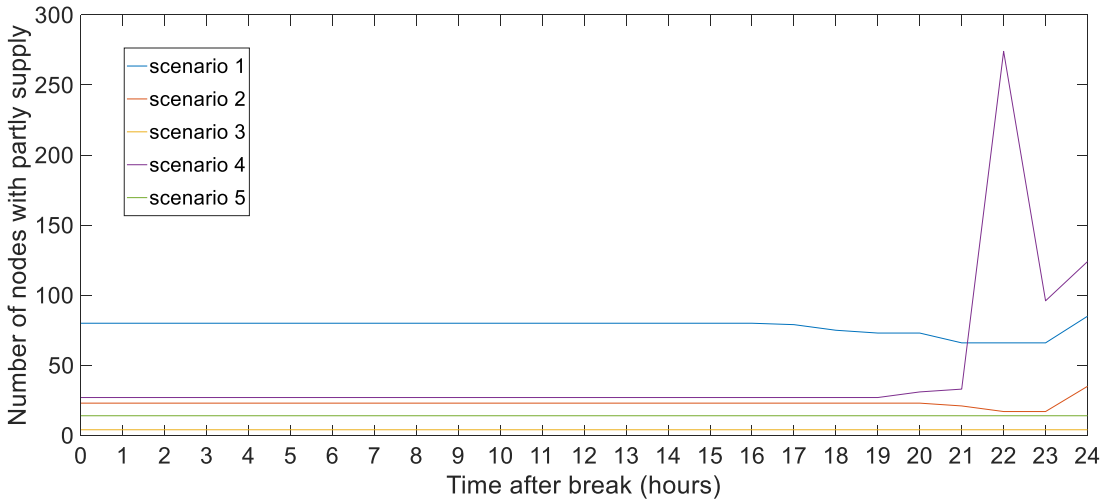

Figure 7.5 Number of nodes with partly supply for 24hrs after pipe failure at 4am

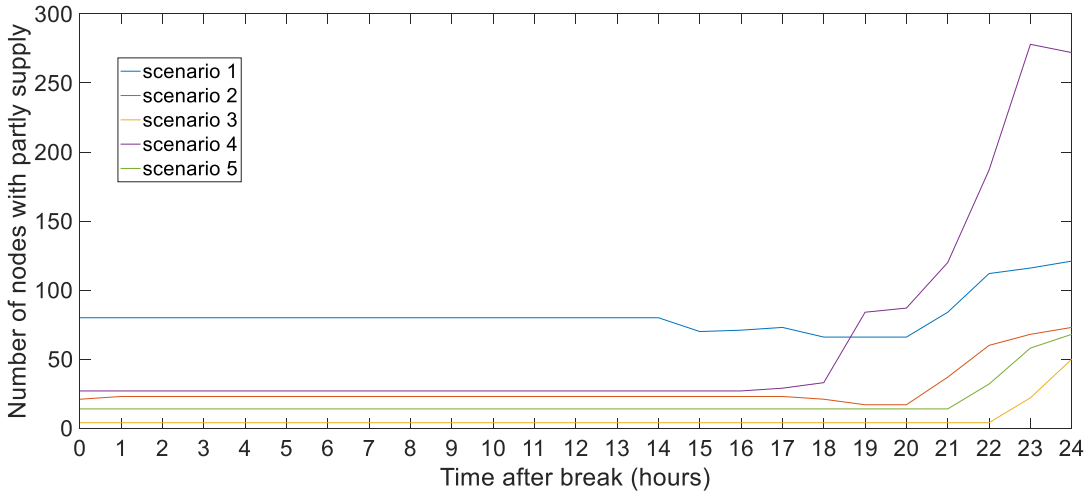

Figure 7.6 Number of nodes with partly supply for 24hrs after pipe failure at 7am

A drawback of the proposed methodology it is assumed that all nodes have an equal importance and the type and vulnerability of the customers fed by each node is ignored because this knowledge is not available. However, in real situations this is not the case e.g. hospitals might be more important, and the consequences are more severe compared to residential area. This shortcoming can be overcome by assigning a weight to each node (Bicik et al. 2009).

Chapter 8. Discussion, Conclusions and Recommendations

8.1 Thesis Summary

The failure of water pipes results in environmental, economic and social costs and, as such, there is growing interest in this area. The overall aim of this proposed methodology is to aid maintenance/rehabilitation/replacement plans and enhance daily allocation and planning of resources.

The literature review presented the factors that contribute to pipe failure. It provided a comprehensive review of the developed methods for pipe failure prediction and impacts assessment. The predictive models were analysed with respect to their suitability for accurately capturing the failure patterns and the impacts approaches with respect to their ability to evaluate the severity of an individual pipe failure.

The failure frequency was predicted long-term, annually and short-term. The long-term and the annual predictions were made using a novel combination of EPR and K-means clustering. The inclusion of a clustering method improves the predictions by using a set of predictive models instead of a single model which captures various failure patterns. A novel method for predicting the annual failure rate considering both weather-related and physical factors was implemented. The *aggregated* EPR models were then used to derive the failure rate of individual pipes. The short-term predictions were based on a novel ANN model that uses the weather data of four days to make predictions for the following two days (non-overlapping). The cross-validation technique was employed to derive a more accurate estimate of the EPR and ANN models' prediction performance and to measure how these results generalize to an independent dataset.

A series of artificial elements was added to all the demand nodes experiencing pressure-deficient condition due to pipe failure, and the satisfied demand was calculated based on the available nodal pressure. The pressure-driven model proposed by Mahmoud et al. (2017) was used in the analysis. The performance indicators employed are the ratio of unsupplied demand and, the percentages of nodes with zero and partly satisfied demands. The outputs of the analysis indicate that failure of most of the pipes results in low impacts. This can be explained by the fact that the examined network is a large real-life network, and a single pipe failure is likely to cause pressure-deficient conditions in a small part of it, whereas performance elsewhere is mostly satisfactory. Also, the redundant design of the WDN allows the flow to reach a given demand node through alternative paths if a single component fails.

8.2 Conclusions and Discussion

In this section, the main points of the methodology and the key findings with respect to the objectives are discussed. Also, the novelty of the thesis is highlighted.

Pipe aggregation

The pipe *aggregation* process is a trade-off between creating groups that are small enough to be uniform and large enough to obtain models with a meaningful 'goodness of fit'. In this thesis, in addition to the pipe attributes (material, diameter, age), the soil type was used as an *aggregation* criterion since soil conditions affect the deterioration rate of the CI pipes. This was also confirmed in the preliminary analysis, which showed that the failure rate can vary substantially depending on the soil type. Also, examining the failure rate within groups of the

same soil type also acts as a surrogate parameter for data that are not easily collected and have inherent uncertainty, such as soil moisture.

K-means clustering method

The burst rate, which is the targeted output, is not used at any stage of clusters' creation since the K-means method would be biased to allocate the homogenous groups with a high failure rate into the same cluster. In that case, the predictions would be accurate for the groups with a high failure rate but of limited usefulness since they constitute only a small part of the network. Furthermore, it would not be feasible to develop a meaningful model for the low failure rate groups due to the limited available data. The K-means method has been previously applied (i.e. Kleiner and Rajani 2012) for creating clusters considering geographical coordinates as criteria, but this implementation with pipe attributes is novel. The K-means method is selected because it is one of the simplest unsupervised learning algorithms that solve the clustering problem without compromising its efficiency.

Cross-validation Technique

The selection of training and test datasets has a significant impact on the accuracy of the models. Therefore, the cross-validation technique was employed to derive a more accurate estimate of model prediction performance by guaranteeing that all observations are used for both training and testing, and that each observation is used for testing exactly once. This is the main advantage of the method compared to others, such as the over-repeated random sub-sampling. Furthermore, it allows one to measure how the results generalize to an independent data set.

Evolutionary Polynomial Regression

Given a set of data, the EPR searches among many possible models to explain those data (Savic et al. 2009). It does, however, require an objective function to ensure the best fit without the introduction of unnecessary complexity (Savic et al. 2009). The key goal, which is to find a systematic way to avoid the problem of over-fitting is addressed by (1) penalizing the complexity of the expression by minimizing the number of terms and; (2) controlling the variance of a_j constants (the variance of estimates) with respect to their values.

Each EPR implementation returns a range of models with a varying number of variables and/or polynomial terms. This is helpful for understanding which inputs are physically meaningful and which can be excluded for the sake of model parsimony while, simultaneously, striving for a degree of generality. The user can assess a set of models looking at different key aspects which encompass the prior knowledge of the phenomenon.

Long-term Predictions

The long-term methodology is novel in combining the K-means method with a data-driven method. It results in a set of EPR models that capture the failure phenomenon in the WDN with a high accuracy. The clustered-based approach captures the variability of failure patterns, particularly the very low and the very high failure rates, better than the single EPR model which captures numerous failure patterns. The number of clusters (i.e. six) is case-specific and relies on data availability. The clustered approach emphasizes on the higher accuracy and does not attempt to find a general optimal number of clusters. Different numbers of clusters were examined until the performance indicators used did not further improve with both the training and test data and there were too few data within

each cluster. Selecting too many clusters could result in several EPR models that need calibration and reduce the available data to an extent that compromises the predictive accuracy of each model (Osman and Bainbridge 2010).

The EPR models calculated the total number of failures as a function of length, diameter and age plus average pressure for the two-polynomial terms model. The substantial difference between the one-polynomial and the two-polynomial models lies on the inclusion of pressure in the list of selected explanatory variables. However, the inclusion of average pressure did not provide a significant increase in the accuracy while indicating a mixed relationship for some clusters (i.e. Clusters 1 and 5). As shown in Figure 6.2, there is not a clear relationship between pressure and failure rate and, therefore, might have limited ability in interpreting the phenomenon. All the homogenous groups entail pipes with the same diameter, age and soil type—contrary to pressure, which is calculated as the ‘’average value’’ of all the pipes. This mixture might have led to the low impact on pipe failure. Other reasons can be that the pressure values are not actual measurements but are the outputs of a calibrated hydraulic EPANET model. The failures have been recorded between 2003 and 2013, whereas the values were taken from EPANET in 2014, and the water utility might have implemented measures to reduce pressure in the meantime.

The two-polynomial terms EPR model did not exhibit a massive improvement in the predictions in terms of the performance indicators, and, for the sake of parsimony and generality while adequately capturing the physical phenomenon, the one-polynomial EPR model was adopted.

A direct relationship was observed between failure and length, while there was an indirect relationship with the diameter—confirming the findings in the literature. The relationship between failure and age is more complex. A direct relationship

was observed for the younger pipes and an inverse relationship for the old pipes, which is explained by the fact that the pipe dataset is left truncated. The age of the oldest pipes is much greater than the time-period in which their failures have been systematically recorded. In addition, the summed length of Clusters 3 and 6, which entail the older pipes, is 15.73% of the network's total length. However, they are 54 out of the 148 homogenous groups, and, therefore, they overwhelm the single-EPR approach. A discrepancy between age and failure frequency has been observed by other researchers, as well (e.g. Boxall et al. 2007; Xu et al. 2011b).

The aggregated EPR models were used to derive the failure rate of individual pipes assuming that all pipes within a group share the same failure rate. This method showed a high accuracy for all the failure rates pointing out the ones with a highest failure propensity. The accurate prediction model is important for the successful implementation of a proactive approach which is part of cost-effective capital investment plans.

Annual Predictions

A distinctive EPR model was developed to examine annual variation of failure rate for each cluster. The direct relationship between failure and FI, which is a surrogate of the severity of low temperatures, indicates that cold spells lead to an increased failure frequency. This finding is confirmed in the preliminary analysis (i.e. Figure 5.6) and attributed to the frost actions, which impose additional load on the buried pipes (Palmer and Williams 2003). The temperature-related candidate explanatory variables were not selected by the one-polynomial term EPR model. Figure 6.10 clearly proved the closer correlation between pipe failure

and FI compared to the rest of candidate explanatory variables. The proposed method was shown to be able to predict peaks of failure frequency (i.e. 2010).

The EPR models calculate the annual number of failures as a function of the FI. This explanatory variable can't be used to directly calculate the number of failures of each homogenous group—as with the long-term predictions and the selected explanatory variables—since it would attribute the same number to all of them. Therefore, the failures were distributed within each cluster's groups proportionally to their length. This approach caused a slight overestimation to the groups with zero or close to zero failures. The inclusion of physical factors improved, overall, the accuracy of the predictions, particularly for the aforementioned groups of pipes. The novelty in the annual predictions is validated by using EPR (i.e. a data-driven method) considering both physical and weather-related factors. The *aggregated* EPR models can be used for identifying the pipes most prone to failure for a specific year.

This approach requires next year's data (which need to be forecasted) to make predictions, introducing a degree of uncertainty. To overcome this uncertainty, the use of lagged explanatory variables—i.e. use of previous and this year's weather data to make predictions for the next year—was attempted. This attempt led to low accuracy since, as indicated in Figure 5.5, the weather data are in line with that specific year's failures. A representative example is Year 2010 when the peak of pipe failure frequency coincides with the greatest FI.

ArcGIS and Jenks Natural Breaks Method

The Jenks natural breaks method is used in conjunction with the ArcGIS mapping tool to visualize the outputs of the EPR models to make them more understandable. ArcGIS is useful for organizing wide ranges of data that are a

basis for decision-making, with applications in similar studies in the water sector (e.g. Burrows et al. 2000; Bicik et al. 2008; Giustolisi and Laucelli 2010; Sitzenfrei et al. 2011; Tabesh and Saber 2012). The Jenks natural breaks method is a popular and efficient method with many applications in various fields for creating choropleth maps, and, due to its advantages, is the default option of ArcGIS. Its main drawback is that it requires the number of desired classes to be indicated before the algorithm is applied to the dataset (North 2009). The selection of the number of classes is not under investigation nor an aim in this study. The banding differs for the long-term and the annual predictions since it is the Jenks method that selects how to best ''spilt up'' the data.

Summary of EPR Predictive Models

Overall, the EPR approach generates accurate results and serves alongside a physically based counterpart to provide a more thorough system characterization. However, when developing failure models for an individual network, it can be questionable if they can be used for generalized asset failure purposes since systems usually differ in one or more significant explanatory variables or, sometimes, even in model structure.

Short-term Predictions

Despite its proven advantages, EPR does not generate satisfactory results for the short-term predictions since the aim here is prediction of failure/no failure and not an actual number (which was the targeted output of the previous EPR models). The initial goal to predict the exact number of failures did not generate a meaningful model with satisfactory accuracy, and, therefore, the aimed output had to be altered. The ANN is employed for the binary model because it is a

powerful data-driven model and has displayed promise for forecasting applications (Zhang et al., 1998; Mounce et al. 2009).

The short-term predictions allow the water utility to enhance their immediate operational planning strategies to accommodate possible increases in pipe failure. The water utility is also aided to meet the standards set by OFWAT to avoid customers' dissatisfaction and comply with the guidelines for water loss.

The input and output are not fixed, and various combinations of them are examined to get the highest correlation since the temperature variation can happen relatively quickly, whereas the potential pipe failure might take longer. The inputs of the model are (1) the minimum air temperature, (2) the maximum air temperature, (3) the mean air temperature, (4) the soil temperature and (5) the FI. Additionally, the output is 1 if there is at least a pipe failure the following days and 0 if not. The selection of failure/not failure as an output is case-specific and relies on the fact that, for most of the days, one or no failures are observed. The approach is applicable to any threshold.

One way to avoid overfitting in ANN models is to use the cross-validation technique (Stone 1974) in which the available data are divided into three sets: training, testing and validation (Shahin et al. 2004). The network is divided into ten equal size folders using the cross-validation method—70% for training, 20% for validation and 10% for testing. The model is fit on the training dataset that is a set of examples used to fit the parameters of the network. The validation dataset provides an unbiased evaluation of a model fit on the training dataset while tuning the model's hyperparameters. It is used for regularization by stopping early—training is stopped when the error on the validation dataset increases, as it is a sign of overfitting. The test dataset assesses the generalization capabilities of the model.

The exhaustive trials led to the conclusion that the use of four consecutive days as input and the following two days as output results in the highest accuracy. The first input is the set of variables for the first four days, and the output is the occurrence of failure(s) in the fifth and sixth day. As indicated by the outputs of the preliminary analysis (Figure 5.6), most of the pipes burst during the coldest months. The cold spells normally last for a couple of days, and there is an increased number of failures during these periods or straight after its end since there might be a time lag between temperature drop and pipe failure. Water pipes burst because the water inside them expands as it gets close to freezing, and this causes an increase in pressure inside the pipe. When the pressure gets too high for the pipe to contain, it ruptures. Another reason for pipe failure can be the freeze and thaw in soil, which causes soil movement. This, in turn, results in a lack of continuous support beneath the pipes, creating bedding stresses.

The proposed method is novel in making short-term predictions since previous approaches (e.g. Rajani and Kleiner 2012; Laucelli et al. 2014) resulted in relationships with low accuracy. It uses recorded data (that do not have to be forecasted) as input variables, which reduces, significantly, the uncertainty in the predictions. The model was successfully applied to a holdout sample, demonstrating that the ANN 'learned' the breakage patterns rather than memorised them.

Pressure-driven Analysis

In the case of a pipe failure, the water flow will change, and the original network will be transformed into a new one with higher internal energy losses, which might make it impossible to deliver the desired flow rate at a minimum delivery pressure (Farmani et al. 2005). Hence, the satisfied demand has been associated with the

nodal pressure changes to assess the performance of the network in a more realistic way (Fujiwara and Ganesharajah 1993). The pipe failure is simulated by adding the following series of artificial elements at all the demand nodes experiencing pressure-deficient conditions (Mahmoud et al. 2017): a Check Valve, an internal dummy Node, a Flow Control Valve and an Emitter. The CV prevents the flow reversal and is the first added artificial element on the demand node. The parameters of the CV are set to produce negligible head losses when water is flowing in the right direction (i.e. short length and large diameter). The downstream dummy node is used just to connect the CV with the TCV since it is an EPANET requirement. The role of the FCV is to ensure that the delivered flow does not exceed the demand at the node. Finally, the emitter is used to represent pressure-dependent demand delivery. The small length, large diameter and large Hazen-Williams coefficient ensure that all additional elements do not introduce significant head loss between a demand node and the emitter.

These artificial elements were selectively added only to pressure deficient demand nodes since the examined network is large and only a small percentage of the nodes is expected to experience pressure drop during a single pipe failure. Also, this thesis examines the failure of distribution and not transmission pipes, which are typically between water treatment works and service reservoirs or between reservoirs and their failure probably affects more nodes. These pressure-deficient nodes are identified by running the DDA-type hydraulic solver (i.e. EPANET) once before the PDA simulation.

Impacts Assessment

Localized asset management decisions enhance the condition of water mains to acceptable level of service since it is impractical and unrealistic due to budget

limitations to replace all the aging pipes simultaneously. An efficient maintenance program should identify the most vulnerable pipes, failure of which could incur significant impacts. This thesis proposes a new combination of a grouped-based method for deriving the failure rate and a pipe-level method for evaluating the impacts on the level of service.

The impacts were evaluated using three performance indicators: the ratio of unsupplied demand, the number of nodes with zero supply and the number of nodes with partial supply. The widely used ratio of unsupplied demand on its own is a rough indicator because, for the same ratio, the number of customers affected can be different. The desired pressure threshold was set to 15m, which is the value above which the WDN operates properly based on the outputs of the calibrated EPANET model, and the minimum pressure was set to 0m. This threshold is in line with OFWAT, which requires low pressure incidents (i.e. drops of pressure below 15m of head) to be reported by every water utility.

The proposed approach considered the geographic location of the failed pipe, the time the failure occurs and its duration. It has been assumed that the failure occurs either when there is a peak in pressure or a peak in demand, and both cases are simulated for an EPS since the failure duration is not known. The diurnal demand variation in nodes, the water level in storage tanks and the valve/pump control settings are considered over this predefined simulation period by changing the parameters of the connected FCVs and emitters according to the current values of the desired demands in deficient nodes.

In both scenarios, the failure of most pipes resulted in low impacts (i.e. low ratio of unsupplied demand and small number of affected nodes) since the examined network is a large real-life network and a failure of a distribution pipe is likely to affect a small part of it, whereas performance elsewhere in the WDN is mostly

satisfactory. Also, the complex structure of WDNs allows them to recover from failures, exploiting the topological redundancy provided by closed loops, so that the flow could reach a given demand node through alternative paths (Tanyimboh et al. 2016; Di Nardo et al.2017b). A redundant network ensures that if a single component fails, there is sufficient residual capacity to provide all flow requirements (Awumah et al. 1991).

Due to the huge number of scenarios (i.e. 18872x2) it is not feasible to present the results of all of them. Therefore, the five pipes (i.e. 5x2 scenarios) that resulted in the highest ratio of unsupplied demand were selected to illustrate the impacts over time. The diameter of these five pipes is 300mm, which is the largest in the examined dataset. Larger diameter pipes tend to have more severe impacts (greater water loss) since they convey more water.

Pipe failure originally leads to uncontrolled flow from the network and, in general, drop in pressure in a number of nodes. Service may be completely lost if the inflow is less than the total outflow (uncontrolled flow plus nodal consumptions), exhausting the reserved supplies (e.g. service reservoir storage). In the examined network, there are eight reservoirs, and, therefore, the increase in unsupplied demand is mostly small. The fluctuation in the number of nodes affected is similar with the ratio of unsupplied demand and remains mostly constant. There is a difference between the numbers of nodes affected totally and partially since, in most scenarios, there is a small drop of pressure (actual pressure is less than the required but still above zero), leading to partially supplied demand. Hence, the number of nodes with zero supply is very low. Overall, the small number of nodes (both zero and partial supply) experiencing pressure-deficient conditions indicates that the surrounding nodes are mainly affected, whereas performance elsewhere in the WDN is mostly satisfactory.

8.3 Recommendations for future work

A few potential topics for future research have been identified following on the work presented in this thesis. Those topics are associated with the pipe failure prediction and the impacts assessment.

1. The accuracy of the predictive models can be further improved by considering factors that are not available in this analysis. These factors include external loads, such as traffic loads, pipe bedding, the implementation of cathodic protection (if any) and the quality of the conveyed water. Furthermore, actual measurements of hydraulic pressure may assist in achieving higher accuracy

2. In this thesis, it is assumed that all nodes have an equal importance, and the type and vulnerability of the customers fed by each node is ignored because this knowledge is not available. However, in real situations, this is not the case, as areas such as hospitals might be more important, and the consequences are more severe compared to residential area. This shortcoming can be overcome by assigning a weight to each node.

3. Pipe failure implies a cost for rehabilitation/replacement and can possibly cause loss of business, costs associated with emergency response and damage to other existing nearby infrastructures. The inclusion of those costs can more accurately illustrate the impacts of pipe failure. Also, the aspect of water quality (e.g. discolouration) should be considered.

Bibliography

Achim, D., F. Ghotb and K. J. McManus (2007). Prediction of water pipe asset life using neural networks. Journal of Infrastructure Systems, 13 (1), 26-30

Ackley, J. R. L., T. T. Tanyimboh, B. Tahar and A. B. Templeman (2001). Head-driven analysis of water distribution systems. Water software systems: theory and applications, 1, 183-192

Agrawal, M. L., R. Gupta and P. R. Bhave (2007). Optimal design of level 1 redundant water distribution networks considering nodal storage. Journal of Environmental Engineering, 133 (3), 319-330

Ahn, J., S. Lee, G. Lee and J. Koo (2005). Predicting water pipe breaks using neural network. Water Science and Technology: Water Supply, 5 (3-4), 159-172

Al-Barqawi, H and T. Zayed (2006). Condition rating model for underground infrastructure sustainable water mains. Journal of Performance of Constructed Facilities, 20 (2), 126-135

Al-Barqawi, H. and T. Zayed. (2008). Infrastructure management: Integrated AHP/ANN model to evaluate municipal water mains' performance. Journal of Infrastructure Systems, 14 (4), 305-318

Alvisi, S. and M. Franchini (2010). Comparative analysis of two probabilistic pipe breakage models applied to a real water distribution system. Civil Engineering and Environmental Systems, 27 (1), 1-22

Ammar, M. A., O. Moselhi and T. M. Zayed (2012). Decision support model for selection of rehabilitation methods of water mains. Structure and Infrastructure Engineering, 8 (9), 847-855

Andreou, S. (1986). Predictive models for pipe break failures and their implications on maintenance planning strategies for deteriorating water distribution systems. PhD thesis, MIT, Cambridge, USA

Andreou, S. A., D. H. Marks and R. M. Clark (1987). A new methodology for modelling break failure patterns in deteriorating water distribution systems: Theory. Advances in Water Resources, 10 (1), 2–10

Ang, W. K. and P. W. Jowitt (2006). Solution for water distribution systems under pressure-deficient conditions. Journal of water resources planning and management, 132 (3), 175-182

Asnaashari, A., E. A. McBean, I. Shahrour and B. Gharabagh (2009). Prediction of water main failure frequencies using multiple and Poisson regression. Water Science and Technology Water Supply, 9 (1), 9-19

Asnaashari, A., E. A. McBean, B. Gharabaghi and D. Tutt (2013). Forecasting watermain failure using artificial neural network modelling. Canadian Water Resources Journal, 38 (1), 24-33

Atkinson, S., R. Farmani, F. A. Memon and D. Butler (2014). Reliability indicators for water distribution system design: Comparison. Journal of Water Resources Planning and Management, 140 (2), 160-168

Awumah, K., I Goulter and S. K. Bhatt (1990). Assessment of reliability in water distribution networks using entropy based measures. Stochastic Hydrology and Hydraulics, 4 (4), 309-320

Awumah, K., I. Goulter and S. K. Bhatt (1991). Entropy-based redundancy measures in water-distribution networks. Journal of Hydraulic Engineering, 117 (5), 595-614

Babayan, A. V., D. A. Savic, G. A., Walters and Z. Kapelan (2007). Robust least-cost design of water distribution networks using redundancy and integration-based methodologies. Journal of Water Resources Planning and Management, 133 (1), 67-77

Babu, K. S. J. and S. Mohan (2012). Extended period simulation for pressure-deficient water distribution network. Journal of Computing in Civil Engineering, 26 (4), 498-505

Baek, C. W., H. D. Jun and J. H. Kim (2010). Development of a PDA model for water distribution systems using harmony search algorithm. KSCE Journal of civil engineering, 14 (4), 613-625

Bakker, M., J. H. G. Vreeburg, L. C. Rietveld and M. Van De Roer (2012). Reducing customer minutes lost by anomaly detection?. *In the Proceedings of the 14th Water Distribution Systems Analysis Conference (WDSA), 24-27 September 2012 in Adelaide, South Australia (p. 913). Engineers Australia*

Baird, G. M. (2011). Who stole my water? The case for water loss control and annual water audits. American Water Works Association. Journal, 103 (10), 22

Beal, C. D. and R. A. Stewart (2013). Identifying residential water end uses underpinning peak day and peak hour demand. Journal of Water Resources Planning and Management, 140 (7), 04014008

Berardi, L., Z. Kapelan, O. Giustolisi and D. A. Savic (2008). Development of pipe deterioration models for water distribution systems using EPR. Journal of Hydroinformatics, 10 (2), 113-126

Berardi, L., R. Ugarelli, J. Røstum and O. Giustolisi, O. (2014). Assessing mechanical vulnerability in water distribution networks under multiple failures. Water Resources Research, 50 (3), 2586-2599

Bhave, P. R. (1981). Node flow analysis distribution systems. Transportation Engineering Journal of ASCE, 107 (4), 457-467

Bicik, J., C. Makropoulos, D. Joksimović, Z. Kapelan, M. S. Morley and D. A. Savic (2008). Conceptual risk-based decision support methodology for improved near real-time response to wds failures. *In the Proceedings of the Water Distribution Systems Analysis (WDSA) 2008 (pp. 1-10)*

Bicik, J., Z. Kapelan and D. A. Savic (2009). Operational perspective of the impact of failures in water distribution systems. *In the Proceedings World Environmental and Water Resources Congress 2009: Great Rivers (pp. 1-10)*

Bicik, J. (2010). A risk-based decision support system for failure management in water distribution networks. PhD Thesis, University of Exeter, UK

Bicik, J., Z. Kapelan, C. Makropoulos and D. A. Savic (2011). Pipe burst diagnostics using evidence theory

BITRE (Bureau of Infrastructure, Transport and Regional Economics). (2014). Yearbook 2014: Australian infrastructure statistical report. Canberra, ACT, Australia

Bonvin, G., A. Samperio, C. Le Pape, V. Mazauric, S. Demassey and N. Maïzi, (2015). A heuristic approach to the water networks pumping scheduling issue. Energy Procedia, 75, 2846-2851

Boxall, J. B., A. O'Hagan, S. Pooladsaz, A. J. Saul and D. M. Unwin (2007). Estimation of burst rates in water distribution mains. Institution of Civil Engineers Water Management, 160 (2), 73–82

Brion, L. M. and L. W. Mays (1991). Methodology for optimal operation of pumping stations in water distribution systems. Journal of Hydraulic Engineering, 117 (11), 1551-1569

Bruaset, S. and S Sægrov (2018). An Analysis of the Potential Impact of Climate Change on the Structural Reliability of Drinking Water Pipes in Cold Climate Regions. Water, 10 (4), 411

Bunn, S. M. and L. Reynolds (2009). The energy-efficiency benefits of pumps-scheduling optimization for potable water supplies. IBM Journal of Research and Development, 53 (3), 388-400

Burns, P., D. Hope and J. Roorda (1999). Managing infrastructure for the next generation. Automation in construction, 8 (6), 689-703

Brewer, C. A. (2006). Basic mapping principles for visualizing cancer data using geographic information systems (GIS). American journal of preventive medicine, 30 (2), 25-36

Burrows, R., G. S. Crowder and J. Zhang, J. (2000). Utilisation of network modelling in the operational management of water distribution systems. Urban Water, 2 (2), 83-95

Canadian Water and Wastewater Association, Municipal Water and Wastewater Infrastructure: Estimated Investment Needs 1997 – 2012, A report partially sponsored by the Canadian Mortgage and Housing Corporation, (1997)

Carrión, A., H. Solano, M. L. Gamiz and A. Debón (2010). Evaluation of the reliability of a water supply network from right-censored and left-truncated break data. Water resources management, 24 (12), 2917-2935

Chandapillai, J. (1991). Realistic simulation of water distribution system. Journal of Transportation Engineering, 117 (2), 258-263

Chik, L., D. Albrecht and J. Kodikara (2016). Estimation of the Short-Term Probability of Failure in Water Mains. Journal of Water Resources Planning and Management, 143 (2), 04016075

Chik, L., D. Albrecht and J. Kodikara (2018). Modeling Failures in Water Mains Using the Minimum Monthly Antecedent Precipitation Index. Journal of Water Resources Planning and Management, 144 (4), 06018004

Cinar, D. and G. Kayakutlu (2010). Scenario analysis using Bayesian networks: A case study in energy sector. Knowledge-Based Systems, 23 (3), 267-276

Clair St, A. M. and S. Sinha (2012). State-of-the-technology review on water pipe condition, deterioration and failure rate prediction models. Urban Water Journal, 9 (2), 85-112

Clair St, A. M. and S. Sinha (2014). Development of a Standard Data Structure for Predicting the Remaining Physical Life and Consequence of Failure of Water Pipes. Journal of Performance of Constructed Facilities, 28 (1), 191-203

Clark, R. M., C. L. Stafford and J. A. Goodrich (1982). Water distribution systems: A spatial and cost evaluation. Journal of the Water Resources Planning and Management Division, 108 (3), 243-256

Clark, R. M., J. Carson, R. C. Thurnau, R. Krishnan and S. Panguluri (2010). Condition assessment modeling for distribution systems using shared frailty analysis. Journal of American Water Works Association, 102 (7), 81-91

Constantine, A. G., J. Darroch and R. Miller (1996). Predicting underground pipeline failure. Water Journal of Australian Water Association 23 (2), 9–10

Cox, D. R. (1992). Regression models and life-tables. In Breakthroughs in statistics (pp. 527-541). Springer New York

Crout, N. M., D. Tarsitano and A. T. Wood (2009). Is my model too complex? Evaluating model formulation using model reduction. Environmental Modelling & Software, 24(1), 1-7

Dandy, G. C. and M. O. Engelhardt (2006). Multi-objective trade-offs between cost and reliability in the replacement of water mains. Journal of Water Resources Planning and Management, 132 (2), 79-88

Davidson, J. W., D. A. Savic and G. A. Walters (1999). Method for the identification of explicit polynomial formulae for the friction in turbulent pipe flow. Journal of Hydroinformatics, 1 (2), 115-126

Davies, C., D. L. Fraser, P. C. Hertzler and R. T. Jones (1997). USEPA's infrastructure needs survey. American Water Works Association. Journal, 89 (12), 30

Davis, P., S. Burn, M. Moglia and S. Gould. (2007). A physical probabilistic model to predict failure rates in buried PVC pipelines. Reliability Engineering & System Safety, 92 (9), 1258-1266

Davis, P., D. De Silva, D. Marlow, M. Moglia, S. Gould and S. Burn (2008). Failure prediction and optimal scheduling of replacements in asbestos cement water pipes. Aqua-Journal of Water Supply, 57 (4), 239-252

D'Ercole, M., M. Righetti, G. S. Raspati, P. Bertola and R. Maria Ugarelli (2018). Rehabilitation Planning of Water Distribution Network through a Reliability—Based Risk Assessment. Water, 10 (3), 277

Debón, A., A. Carrión, E. Cabrera and H. Solano (2010). Comparing risk of failure models in water supply networks using ROC curves. Reliability Engineering and System Safety, 95 (1), 43-48

Dehghan, A., K. J. McManus and E. F. Gad (2009). Probabilistic failure prediction for deteriorating pipelines: Nonparametric approach. Journal of Performance of Constructed Facilities, 22 (1), 45-53

Demissie, G., S. Tesfamariam and R. Sadiq (2017). Prediction of Pipe Failure by Considering Time-Dependent Factors: Dynamic Bayesian Belief Network Model. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 3 (4), 04017017

Diao, K., C. Sweetapple, R. Farmani, G. Fu, S. Ward and D. Butler (2016). Global resilience analysis of water distribution systems. Water research, 106, 383-393

Díaz, S., R. Mínguez and J. González, J. (2016). Stochastic approach to observability analysis in water networks. Ingeniería del Agua, 20 (3), 139-152

Di Nardo, A., M. Di Natale, C. Giudicianni, D. Musmarra, J. R. Varela, G. F. Santonastaso, A. Simone and V. Tzatchkov (2017a). Redundancy features of water distribution systems. Procedia Engineering, 186, 412-419

Di Nardo, A., M. Di Natale, C. Giudicianni, G. F Santonastaso and D. A. Savic (2017b). Simplified approach to water distribution system management via identification of a primary network. Journal of Water Resources Planning and Management, 144 (2), 04017089

Doyle, G., M. V. Seica and M. W.Grabinsky. (2003). The role of soil in the external corrosion of cast iron water mains in Toronto, Canada. Canadian geotechnical journal, 40 (2), 225-236

Dridi, L., A. Mailhot, M. Parizeau and J. P. Villeneuve (2009). Multiobjective approach for pipe replacement based on bayesian inference of break model parameters. Journal of Water Resources Planning and Management, 135 (5), 344-354

Duncan, A., D. Tyrrell, N. Smart, E. Keedwell, S. Djordjevic and D. A. Savic (2013). Comparison of machine learning classifier models for bathing water quality exceedances in UK. *In Proceedings of the 35th International Association for Hydro-Environment Engineering and Research Congress (IAHR 2013)*, Chengdu, China

Economou, T., Z. Kapelan and T. C. Bailey (2007). An aggregated hierarchical Bayesian model for the prediction of pipe failures. *In Proceedings of 9th International Conference on Computing and Control for the Water Industry (CCWI)*, Leicester, UK

Economou, T., Z. Kapelan and T. C. Bailey (2008). A zero-inflated Bayesian model for the prediction of water pipe bursts. *In Proceedings of the 10th Annual Water Distribution Systems Analysis Conference (WDSA)*, Kruger National Park, South Africa

Economou, T., Z. Kapelan and T. C. Bailey (2012). On the prediction of underground water pipe failures: zero inflation and pipe-specific effects. Journal of Hydroinformatics, 14 (4), 872-883

Eghbali, A. H., K. Behzadian, F. Hooshyaripor, R. Farmani and A. P. Duncan (2017). Improving prediction of dam failure peak outflow using neuroevolution combined with k-means clustering. Journal of Hydrologic Engineering, 22 (6), 04017007

El-Baroudy, I., A. Elshorbagy, S. K. Carey, O. Giustolisi and D. A. Savic (2010). Comparison of three data-driven techniques in modelling the evapotranspiration process. Journal of Hydroinformatics, 12 (4), 365-379

Engelhardt, M. O., P.J. Skipworth, D. A. Savic, A. J. Saul and G. A. Walters (2000). Rehabilitation strategies for water distribution networks: a literature review with a UK perspective. Urban Water, 2 (2), 153-170

Engelhardt, M., D. A. Savic, P. Skipworth, A. Cashman, A. Saul, and G. A. Walters (2003). Whole life costing: Application to water distribution network. Water Science and Technology: Water Supply 3 (1-2), 87-93

Fahmy, M. and O. Moselhi (2009). Forecasting the remaining useful life of cast iron water mains. Journal of performance of constructed facilities, 23 (4), 269-275

Farley, M., and S. Trow (2003). Losses in water distribution networks. IWA publishing

Farmani, R., G. A. Walters and D. A. Savic (2005). Trade-off between total cost and reliability for Anytown water distribution network. Journal of Water Resources Planning and Management, 131 (3), 161-171

Fayyad, U., G. Piatetsky-Shapiroand and P. Smyth (1996). From data mining to knowledge discovery in databases. Al magazine, 17 (3), 37-54

Filion, Y. R., B. J. Adams and B. W. Karney (2007). Stochastic design of water distribution systems with expected annual damages. Journal of Water Resources Planning and Management, 133 (3), 244-252

Fiore, A., L. Berardi and G. C. Marano (2012). Predicting torsional strength of RC beams by using evolutionary polynomial regression. Advances in Engineering Software, 47 (1), 178-187

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

Francis, R. A., S. D. Guikema and L. Henneman (2014). Bayesian belief networks for predicting drinking water distribution system pipe breaks. Reliability Engineering & System Safety, 130, 1-11

Fuch-Hanusch, D. F., F. Friedl, R. Scheucher, B. Kogseder and D. Muschalla (2013). Effect of seasonal climatic variance on water main failure frequencies in moderate climate regions. Water Science and Technology: Water Supply, 13 (2), 435-446

Fujiwara, O., and T. Ganesharajah (1993). Reliability assessment of water supply systems with storage and distribution networks. Water Resources Research, 29 (8), 2917-2924

Gandhi, T., B. K. Panigrahi and S. Anand (2011). A comparative study of wavelet families for EEG signal classification. Neurocomputing, 74 (17), 3051-3057

Geisser, S., 1975. The predictive sample reuse method with applications. Journal of American Statistics Association, 70, 320–328

Germanopoulos, G. (1985). A technical note on the inclusion of pressure dependent demand and leakage terms in water supply network models. Civil Engineering Systems, 2 (3), 171-179

Germanopoulos, G., P. W. Jowitt and J. P. Lumbers (1986). Assessing the reliability supply and level of service for water distribution systems. *In the Proceedings of the Institution of Civil Engineers*, 80 (2), 413-428

Ghorbanian, V., Y. Guo and B. Karney (2016). Field Data–Based Methodology for Estimating the Expected Pipe Break Rates of Water Distribution Systems. Journal of Water Resources Planning and Management, 142 (10), 04016040

Giudicianni, C., A. Di Nardo, M. Di Natale, R. Greco, G. F. Santonastaso and A. Scala (2018). Topological taxonomy of water distribution networks. Water, 10 (4), 444

Giustolisi, O. (2004). Using genetic programming to determine Chezy resistance coefficient in corrugated channels. Journal of Hydroinformatics, 6 (3), 157-173

Giustolisi, O. and D. Laucelli. (2005). Increasing Generalisation of Input-Output Artificial Neural Networks in Rainfall-Runoff Modelling. Hydrological Sciences Journal, 50 (3), 439-457

Giustolisi, O. and D. A. Savic (2006). A symbolic data-driven technique based on evolutionary polynomial regression. Journal of Hydroinformatics, 8 (3), 207-222

Giustolisi, O., D. Laucelli and D. A. Savic (2006). Development of rehabilitation plans for water mains replacement considering risk and cost-benefit assessment. Civil Engineering and Environmental Systems, 23 (3), 175-190

Giustolisi, O., A. Doglioni, D. A. Savic and B. W. Webb (2007). A multi-model approach to analysis of environmental phenomena. Journal of Environmental Modelling and Software, 22 (5), 674-682

Giustolisi, O., Z. Kapelan and D. A. Savic (2008a). Algorithm for automatic detection of topological changes in water distribution networks. Journal of Hydraulic Engineering, 134 (4), 435-446

Giustolisi, O., Z. Kapelan and D. A. Savic (2008b). Extended period simulation analysis considering valve shutdowns. Journal of water resources planning and management, 134 (6), 527-537

Giustolisi, O., and D. A. Savic (2009). Advances in data-driven analyses and modelling using EPR-MOGA. Journal of Hydroinformatics, 11 (3-4), 225-236

Giustolisi, O., D. A. Savic and D. Laucelli (2009). Asset deterioration analysis using multi-utility data and multi-objective data mining. Journal of Hydroinformatics, 11 (3-4), 211–224

Giustolisi, O. and D. Laucelli (2010). Water distribution network pressure-driven analysis using the enhanced global gradient algorithm (EGGA). Journal of Water Resources Planning and Management, 137 (6), 498-510

Goldberg, D. E. (1989). Genetic Algorithms in Search, Optimization and Machine Learning. Addison Wesley

Gómez-Martínez, P., F. Cubillo, F. J. Martín-Carrasco and L. Garrote (2017). Statistical dependence of pipe breaks on explanatory variables. Water, 9 (3), 158

Gorev, N. B. and I. F. Kodzhespirova (2013). Noniterative implementation of pressure-dependent demands using the hydraulic analysis engine of EPANET 2. Water resources management, 27 (10), 3623-3630

Gould, S. J. F., J. Kodikara, P. Rajeev, X. L. Zhao and S. Burn (2011a). A void ratio–water content–net stress model for environmentally stabilized expansive soils. Canadian Geotechnical Journal, 48 (6), 867-877

Gould, S. J. F., F. A. Boulaire, S. Burn, X. L. Zhao and J. K. Kodikara (2011b). Seasonal factors influencing the failure of buried water reticulation pipes. Water Science and Technology, 63 (11), 2692–2699

Goulter, I. C. and A. Kazemi (1988). Spatial and temporal groupings of water main pipe breakage in Winnipeg. Canadian Journal of Civil Engineering, 15 (1), 91–97

Goulter, I., M. Thomas, L. Mays, B. Sakarya, F. Bouchart and Y.Tung (1999). “Reliability analysis for design.” Water distribution systems handbook, L. W. Mays, ed., McGraw-Hill, New York, 18.1–18.52

Greyvenstein, B. and J. E. Van Zyl (2007). An experimental investigation into the pressure-leakage relationship of some failed water pipes. Journal of Water Supply: Research and Technology-AQUA, 56 (2), 117-124

Grigg, N. S. (2013). Water main breaks: risk assessment and investment strategies. Journal of Pipeline Systems Engineering and Practice, 4 (4), 04013001

Gupta, R. and P. R. Bhave (1994). Reliability analysis of water-distribution systems. Journal of Environmental Engineering, 120 (2), 447-461

Gupta, R. and P. R. Bhave (1996). Comparison of methods for predicting deficient-network performance. Journal of Water Resources Planning and Management, 122 (3), 214-217

Gupta, R., A. Sayyed and M. A. Haider (2013). Predicting deficient condition performance of water distribution networks. Civil Engineering Infrastructures Journal, 46 (2), 161-173

Gupta, R., N. Kakwani and L. Ormsbee (2014). Optimal upgrading of water distribution network redundancy. Journal of Water Resources Planning and Management, 141 (1), 04014043

Habibian, A. (1994). Effect of temperature changes on water-main breaks. Journal of transportation engineering, 120 (2), 312-321

Haider, H., R. Sadiq and S. Tesfamariam (2013). Performance indicators for small-and medium-sized water supply systems: a review. Environmental reviews, 22 (1), 1-40

Herrera, M., E. Abraham and I. Stoianov (2016). A graph-theoretic framework for assessing the resilience of sectorised water distribution networks. Water Resources Management, 30 (5), 1685-1699

Herstein, L. M., Y. R. Filion, and K. R. Hall (2010). Evaluating the environmental impacts of water distribution systems by using EIO-LCA-based multiobjective optimization. Journal of Water Resources Planning and Management 137 (2), 162-172

Herz, R. K. (1998). Exploring rehabilitation needs and strategies for water distribution networks. Journal of Water Supply: Research and Technology-Aqua, 47 (6), 275-283

Holt, J. B., C. P. Lo, and T. W. Hodler (2004). Dasymetric estimation of population density and areal interpolation of census data. Cartography and Geographic Information Science, 31 (2), 103-121

Hu Y and D. W. Hubble (2007). Factors contributing to the failure of asbestos cement water mains. Canadian Journal of Civil Engineer, 34 (5), 608–621

Hudak, P. F., B. Sadler and B. A. Hunter (1998). Analysing underground water-pipe breaks in residual soils. Water Engineering and Management, 145 (12), 15-20

InfraGuide, N. G. (2006). Managing Risk. Decision Making and Investment Planning

Jacobs, H. E. and J. L. Strijdom (2009). Evaluation of minimum residual pressure as design criterion for South African water distribution systems. Water SA, 35 (2), 183-191

Jafar, R., I. Shahrour and L. Juran (2010). Application of Artificial Neural Networks (ANN) to model the failure of urban water mains. Journal of Mathematical and Computer Modelling, 51 (9), 1170-1180

James, G., D. Witten, T. Hastie and R. Tibshirani (2013). An introduction to statistical learning (Vol. 112). New York: springer.

Jang, J. S. R., C. T. Sun and E. Mizutani (1997). Neuro-fuzzy and soft computing; a computational approach to learning and machine intelligence. Practice Hall

Javadi, A. A., M. Rezania and M. M. Nezhad. (2006). Evaluation of liquefaction induced lateral displacements using genetic programming. Computers and Geotechnics, 33 (4), 222-233

Jenks, G. F. (1963). Generalization in statistical mapping. Annals of the Association of American Geographers, 53 (1), 15-26

Jenks, G. F. (1967). The data model concept in statistical mapping. International yearbook of cartography, 7, 186-190

Jenkins, L., S. Gokhale and M. McDonald (2014). Comparison of Pipeline Failure Prediction Models for Water Distribution Networks with Uncertain and Limited Data. Journal of Pipeline Systems Engineering and Practice, 6 (2), 04014012

Ji, J., D. J. Robert, C. Zhang, D. Zhang and J. Kodikara (2017). Probabilistic physical modelling of corroded cast iron pipes for lifetime prediction. Structural Safety, 64, 62-75

Jowitt, P. W. and C. Xu (1993). Predicting pipe failure effects in water distribution networks. Journal of Water Resources Planning and Management, 119 (1), 18-31

Jun, H., G. V. Loganathan, J. H. Kim and S. Park (2008). Identifying pipes and valves of high importance for efficient operation and maintenance of water distribution systems. Water resources management, 22 (6), 719-736

Jun, L. and Y. Guoping (2012). Iterative methodology of pressure-dependent demand based on EPANET for pressure-deficient water distribution analysis. Journal of Water Resources Planning and Management, 139 (1), 34-44

Kabir, G., S. Tesfamariam, A. Francisque and R. Sadiq (2015a). Evaluating risk of water mains failure using a Bayesian belief network model. European Journal of Operational Research, 240 (1), 220-234

Kabir, G., G. Demissie, R. Sadiqand and S. Tesfamariam (2015b). Integrating failure prediction models for water mains: Bayesian belief network-based data fusion. Knowledge-Based Systems, 85, 159-169

Kabir, G., S. Tesfamariam, J. Loeppky and R. Sadiq (2015c). Integrating Bayesian Linear Regression with Ordered Weighted Averaging: Uncertainty

Analysis for Predicting Water Main Failures. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 1 (3), 04015007

Kabir, G., S. Tesfamariam and R. Sadiq (2015d). Predicting water main failures using Bayesian model averaging and survival modelling approach. Reliability Engineering and System Safety, 142, 498-514

Kabir, G., S. Tesfamariam, J. Loeppky and R. Sadiq (2016). Predicting water main failures: A Bayesian model updating approach. Knowledge-Based Systems, 110, 144-156

Kalungi, P. and T. T. Tanyimboh (2003). Redundancy model for water distribution systems. Reliability Engineering & System Safety, 82 (3), 275-286

Kanungo, T., D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman and A. Y. Wu (2002). An efficient k-means clustering algorithm: Analysis and implementation. Pattern Analysis and Machine Intelligence IEEE, 24 (7), 881-892

Kapelan, Z., O. Giustolisi and D. A. Savic (2006a). Risk assessment of water supply interruptions due to mechanical pipe failures. *In Proceedings of the Combined International Conference of Computing and Control for the Water Industry, (CCWI) 2007 and Sustainable Urban Water Management, (SUWM) 2007, (pp. 1-5)*

Kapelan, Z., D. A. Savic, G. A. Walters and A. V. Babayan (2006b). Risk-and robustness-based solutions to a multi-objective water distribution system rehabilitation problem under uncertainty. Water science and technology, 53 (1), 61-75

Kettler, A. J. and I. C. Goulter (1985). An analysis of pipe breakage in urban water distribution networks. Canadian Journal of Civil Engineering 12 (2), 286-293

Kim, J. H., C. W. Baek, D. J. Jo, E. S. Kim and M. J. Park (2004). Optimal planning model for rehabilitation of water networks. Water Science and Technology: Water Supply 4 (3), 133-148

Kim, S. E. and I. W. Seo (2015). Artificial neural network ensemble modeling with conjunctive data clustering for water quality prediction in rivers. Journal of Hydro-Environment Research, 9 (3), 325-339

Kimutai, E., G. Betrie, R. Brander, R. Sadiq and S. Tesfamariam (2015). Comparison of Statistical Models for Predicting Pipe Failures: Illustrative Example with the City of Calgary Water Main Failure. Journal of Pipeline Systems Engineering and Practice

Kleiner, Y., B. J. Adams and J. S. Rogers (1998). Long-term planning methodology for water distribution system rehabilitation. Water resources research, 34 (8), 2039-2051

Kleiner, Y. and B. Rajani (1999). Using limited data to assess future needs. American Water Works Association. Journal, 91 (7), 47

Kleiner, Y. and B. Rajani (2000). Considering time-dependent factors in the statistical prediction of water main breaks. *In the Proceedings of American Water Works Association Infrastructure Conference (AWWA 2000),* Seattle, USA, pp. 1-12

Kleiner, Y. and B. Rajani (2001). Comprehensive review of structural deterioration of water mains: statistical models. Urban water 3 (3), 131-150

Kleiner, Y., B. J. Adams and J. S. Rogers (2001). Water distribution network renewal planning. Journal of Computing in Civil Engineering, 15 (1), 15-26

Kleiner, Y. and B. Rajani (2002). Forecasting variations and trends in water-main breaks. Journal of Infrastructure Systems 8 (4), 122-131

Kleiner, Y., S. McDonald, B. Rajani, Y. Kleiner and S. Mcdonald (2003). Cathodic protection of water mains in Ottawa: analysis and planning. Corrosion control for enhanced reliability and safety, 1-14

Kleiner, Y. and B. Rajani (2004). Quantifying effectiveness of cathodic protection in water mains: theory. Journal of infrastructure systems 10 (2), 43-51

Kleiner, Y. and B. Rajani (2010). Dynamic influences on the deterioration rates of individual water mains (I-WARP). Water Research Foundation Research Report, Denver, USA

Kleiner, Y. and B. Rajani (2012). Comparison of four models to rank failure likelihood of individual pipes. Journal of Hydroinformatics 14 (3), 659-681

Kloog, I., A. Haim, R. G. Stevens, M. Barchana and B. A. Portnov (2008). Light at night co-distributes with incident breast but not lung cancer in the female population of Israel. Chronobiology international, 25 (1), 65-81

Kohavi, R., (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *In Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI 1995)*, 2 (12), 1137-1143

Kohavi, R. and F. Provost (1998). Glossary of terms. Machine Learning 30 (2-3), 271-274

Kovalenko, Y., N. B. Gorev, I. F. Kodzhespirova, E. Prokhorov and G. Trapaga. (2014). Convergence of a hydraulic solver with pressure-dependent demands. Water resources management, 28 (4), 1013-1031

Kumar, R. and A. Indrayan (2011). Receiver operating characteristic (ROC) curve for medical researchers. Indian pediatrics, 48 (4), 277-287

Kunkel, G., K. Laven and B. Mergelas (2008). Does Your City Have High-risk Pipes?. Journal-American Water Works Association, 100 (4), 70-74

Kutyłowska, M. and H. Hotloś (2014). Failure analysis of water supply system in the Polish city of Głogów. Engineering Failure Analysis, 41, 23-29

Large, A., Y. L. Gat, S. M. Elachachi, E. Renaud and D. Breysse (2015). Decision support tools: Review of risk models in drinking water network asset management. Water Utility Journal, 10, 45–53

Laucelli, D., B. Rajani, Y. Kleiner and O. Giustolisi (2014). Study on relationships between climate-related covariates and pipe bursts using evolutionary-based modelling. Journal of Hydroinformatics, 16 (4), 743-757

Le Gat, Y. (2009). Une extension du processus de Yule pour la modélisation stochastique des événements récurrents. Application aux défaillances de canalisations d'eau sous pression. PhD Thesis, Cemagref Bordeaux, Paristech, France

Le Gat, Y. and P. Eisenbeis (2000). Using maintenance records to forecast failures in water networks. Urban Water 2 (3), 173-181

Lei, J. and S. Sægrov (1998). Statistical approach for describing failures and lifetimes of water mains. Water Science and Technology, 38 (6), 209-217

Li, C. Q. and M. Mahmoodian (2013). Risk based service life prediction of underground cast iron pipes subjected to corrosion. Reliability Engineering and System Safety, 119, 102-108

Li, Z., B. Zhang, Y. Wang, F. Chen, R. Taib, V. Whiffin and Y. Wang (2014). Water pipe condition assessment: a hierarchical beta process approach for sparse incident data. Machine learning, 95 (1), 11-26

Li, F., L. Ma, Y. Sun and J. Mathew (2015). Optimized group replacement scheduling for water pipeline network. Journal of Water Resources Planning and Management, 142 (1), 04015035

Lim, S. R., D. Park, D. and J. M. Park (2008). Analysis of effects of an objective function on environmental and economic performance of a water network system using life cycle assessment and life cycle costing methods. Chemical Engineering Journal, 144 (3), 368-378

Lindhe, A., L. Rosén, T. Norberg and O. Bergstedt (2009). Fault tree analysis for integrated and probabilistic risk analysis of drinking water systems. Water research, 43 (6), 1641-1653

Liserra, T., M. Maglionico, V. Ciriello and V. Di Federico (2014). Evaluation of reliability indicators for WDNs with demand-driven and pressure-driven models. Water resources management, 28 (5), 1201-1217

Liu, Z., R. Sadiq, B. Rajani and H. Najjaran (2009). Exploring the relationship between soil properties and deterioration of metallic pipes using predictive data mining methods. Journal of Computing in Civil Engineering, 24 (3), 289-301

Loganathan, G. V., S. Park and H. D. Sherali (2002). Threshold break rate for pipeline replacement in water distribution systems. Journal of water resources planning and management 128 (4), 271-279

Ma, S., J. Li, L. Liu and T. D. Le (2016). Mining combined causes in large data sets. Knowledge-Based Systems, 92, 104-111

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *In: Proceedings of the 5th Berkeley symposium on mathematical statistics and probability, 1 (14), 281-297*

Mahmoud, H. A., D. A. Savic and Z. Kapelan (2017). New pressure-driven approach for modeling water distribution networks. Journal of Water Resources Planning and Management, 143 (8), 04017031

Mailhot, A., G. Pelletier, J. F. Noël and J. P. Villeneuve (2000). Modeling the evolution of the structural state of water pipe networks with brief recorded pipe break histories: Methodology and application. Water Resources Research 36 (10), 3053-3062

Makar, J. M. (2000). A preliminary analysis of failures in grey cast iron water pipes. Engineering Failure Analysis 7 (1), 43-53

Makar, J. M. and Y. Kleiner (2000). Maintaining water pipeline integrity. In Proceedings of American Water Works Association Infrastructure Conference and Exhibition, Baltimore

Malm A., O. Ljunggren, O. Bergstedt, T. J. Petterssonand and G. M. Morrison (2012). Replacement predictions for drinking water networks through historical data. Water research, 46 (7), 2149-2158

Mansoor, M. A. M. (2007) Performance assessment of water distribution systems. PhD Thesis, University of Loughborough, UK

Marlow, D., S. Gould, D. Beal and B. Lane (2015). Rehabilitation of Small-Diameter Cast-Iron Pipe: US, UK, and Australian Perspectives. Journal-American Water Works Association, 107 (1), 12-21

Martínez-Codina, Á., M. Castillo, D. González-Zeasand L. Garrote (2015a). Pressure as a predictor of occurrence of pipe breaks in water distribution networks. Urban Water Journal, 1-11

Martínez-Codina, Á., L. Cueto-Felgueroso, M. Castillo and L. Garrote (2015b). Use of pressure management to reduce the probability of pipe breaks: A Bayesian approach. Journal of Water Resources Planning and Management, 141 (9), 04015010

Martins, A., J. P. Leitão and C. Amado (2013). Comparative Study of Three Stochastic Models for Prediction of Pipe Failures in Water Supply Systems. Journal of Infrastructure Systems 19 (4), 442-450

McMaster, R. (1997). In Memoriam: George F. Jenks (1916-1996). Cartography and Geographic Information Systems, 24 (1), 56-59

McNeill, L. S. and M. Edwards (2002). The importance of temperature in assessing iron pipe corrosion in water distribution systems. Environmental Monitoring and Assessment, 77 (3), 229-242

Meng, F., G. Fu, R. Farmani, C. Sweetapple and D. Butler (2018). Topological attributes of network resilience: A study in water distribution systems. Water research, 143, 376-386

Michaud, D. and G. E. Apostolakis (2006). Methodology for ranking the elements of water-supply networks. Journal of infrastructure systems, 12 (4), 230-242

Monie, W. D. and C. M. Clark (1974). Loads on underground pipe due to frost penetration. Journal of American Water Works Association, 353-358

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel and T. L. Veith (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the Asabe, 50 (3), 885-900

Morris, R. E. (1967). Principal causes and remedies of water main breaks. Journal of American Water Works Association, 59 (7), 782-798

Moselhi, O. and M. Fahmy (2007). Discussion of 'Prediction of water pipe asset life using neural networks' by D. Achim, F. Ghotb, and K. J. McManus. Journal of Infrastructure System, 14 (3), 272–273

Mounce, S. R., J. B. Boxall and J. Machell, J. (2009). Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows. Journal of Water Resources Planning and Management, 136 (3), 309-318

Mounce, S. R. and J. B. Boxall (2010). Implementation of an on-line artificial intelligence district meter area flow meter data analysis system for abnormality detection: a case study. Water Science & Technology: Water Supply, 10 (3), 437–444

Mounce, S. R., K. Ellis, J. M. Edwards, V. L. Speight, N. Jakomis and J. B. Boxall (2017). Ensemble decision tree models using rusboost for estimating risk of iron failure in drinking water distribution systems. Water Resources Management, 31 (5), 1575-1589

Muhammed, K., R. Farmani, K., Behzadian, K., Diao and D. Butler (2017). Optimal rehabilitation of water distribution systems using a cluster-based technique. Journal of Water Resources Planning and Management, 143 (7), 04017022

Muhlbauer, W. K. (2004). Pipeline risk management manual: ideas, techniques, and resources. Gulf Professional Publishing

Nafi, A. and Y. Kleiner (2010). Scheduling renewal of water pipes while considering adjacency of infrastructure works and economies of scale. Journal of water resources planning and management, 136 (5), 519-530

Newport, R. (1981). Factors influencing the occurrence of bursts in iron water mains. Aqua, (3), 274-278

Nishiyama, M. and Y. Filion (2013). Review of Statistical Water Main Break Prediction Models: 2002-2012. Canadian Journal of Civil Engineering, 40 (10), 972-979

Nishiyama, M. and Y. Filion (2014). Forecasting breaks in cast iron water mains in the city of Kingston with an artificial neural network model. Canadian Journal of Civil Engineering, 41 (10), 918-923

North, M. A. (2009). A method for implementing a statistically significant number of data classes in the Jenks algorithm. *In the proceeding of the Sixth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2009).* IEEE, pp. 35–38

O'Day, D. K. (1987). Water distribution record keeping and planning approaches: current utility practices. Journal of the New England Water Works Association JNEWA 6, 101 (2), 145-168

OFWAT. (2008). June return 2008 reporting requirements. Retrieved from: http://www.ofwat.gov.uk/regulating/junereturn/prs_web_jrreportingreq

Ormsbee, L. and A. Kessler (1990). Optimal upgrading of hydraulic-network reliability. Journal of Water Resources Planning and Management, 116 (6), 784-802

Osman, H. and K. Bainbridge (2010). Comparison of statistical deterioration models for water distribution networks. Journal of Performance of Constructed Facilities, 25 (3), 259-266

Ostfeld, A., D. Kogan and U. Shamir (2002). Reliability simulation of water distribution systems–single and multiquality. Urban Water, 4 (1), 53-61

Ozger, S. S. (2003). A semi-pressure-driven approach to reliability assessment of water distribution networks. PhD Thesis, Arizona State University, USA

Pacchin, E., S. Alvisi and M. Franchini (2017). A new non-iterative method for pressure-driven snapshot simulations with EPANET. Procedia Engineering, 186, 135-142

Paez, D., C. R. Suribabu and Y. Filion (2018). Method for Extended Period Simulation of Water Distribution Networks with Pressure Driven Demands. Water Resources Management, 32 (8), 2837-2846

Palmer, A. C. and P. J. Williams (2003). Frost heave and pipeline upheaval buckling. Canadian Geotechnical Journal, 40 (5), 1033-1038

Park, S., H. Jun, B. J. Kim and G. C. Im (2008). Modelling of water main failure rates using the log-linear ROCOF and the power law process. Water Resources Management, 22 (9), 1311-1324

Park, S., H. Jun, N. Agbenowosi, B. J. Kim and K. Lim (2011). The proportional hazards modelling of water main failure data incorporating the time-dependent effects of covariates. Water Resources Management, 25 (1), 1-19

Pearl, J. (2014). Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann

Pelletier, G., A. Mailhot and J. P. Villeneuve (2003). Modeling water pipe breaks-three case studies. Journal of Water Resources Planning and Management, 129 (2), 115-123

Pham, D. T., S. S. Dimov and C. D. Nguyen (2005). Selection of K in K-means clustering. In Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 219 (1), 103-119

Phan, H. C., A. S. Dhar and R. Sadiq (2018). Prioritizing Water Mains for Inspection and Maintenance Considering System Reliability and Risk. Journal of Pipeline Systems Engineering and Practice, 9 (3), 04018009

Pietrucha-Urbanik, K., and A. Studziński (2018). Qualitative analysis of the failure risk of water pipes in terms of water supply safety. Engineering Failure Analysis, 95, 371-378

Prechelt, L. (1998). Early stopping-but when?. In Neural Networks: Tricks of the trade (pp. 55-69). Springer, Berlin, Heidelberg.

Rahmani, F., K. Behzadian and A. Ardeshir (2015). Rehabilitation of a water distribution system using sequential multiobjective optimization models. Journal of Water Resources Planning and Management, 142 (5), C4015003

Rajani, B. and C. Zhan (1996). On the estimation of frost loads. Canadian geotechnical journal, 33 (4), 629-641

Rajani, B., C. Zhan and S. Kuraoka. (1996). Pipe–soil interaction analysis of jointed water mains. Canadian Geotechnical Journal, 33 (3), 393–404

Rajani, B. and Y. Kleiner (2001). Comprehensive review of structural deterioration of water mains: physically based models. Urban water, 3 (3), 151-164

Rajani, B. and Y. Kleiner (2002). Towards pro-active rehabilitation planning of water supply systems. *In the Proceedings of the International Conference on Computer Rehabilitation of Water Networks-CARE-W*, Dresden, Germany

Rajani, B. and Y. Kleiner (2003). Protection of Ductile Iron Water Mains against External Corrosion: Review of Methods and Case Histories. Journal American Water Works Association, 95 (11), 110-125

Rajani, B. and S. Tesfamariam (2004). Uncoupled axial, flexural, and circumferential pipe soil interaction analyses of partially supported jointed water mains. Canadian geotechnical journal, 41 (6), 997-1010

Rajani, B. and Y. Kleiner (2007). Quantifying effectiveness of cathodic protection in water mains: Case studies. Journal of Infrastructure Systems, 13 (1), 1-11

Rajani, B., Y. Kleiner and J. Sink (2012). Exploration of the relationship between water main breaks and temperature covariates. Urban Water Journal 9 (2), 67-84

Rajeev, P., D. Chan and J. Kodikara (2012). Ground–atmosphere interaction modelling for long-term prediction of soil moisture and temperature. Canadian Geotechnical Journal, 49 (9), 1059-1073

Rajeev, P., J. Kodikara, D. Robert, P. Zeman and B. Rajani (2014). Factors contributing to large diameter water pipe failure. Water Asset Management International, 10 (3), 9-14

Reddy, L. S. and K. Elango (1989). Analysis of water distribution networks with head-dependent outlets. Civil Engineering Systems, 6 (3), 102-110

Redmond, S. J. and C. Heneghan (2007). A method for initialising the K-means clustering algorithm using kd-trees. Pattern recognition letters, 28 (8), 965-973

Reed, C., A. Robinson and D. Smart. (2007). Potential techniques for the assessment of joints in water distribution pipelines. American Water Works Research Foundation

Reddy, L. S. and K. Elango (1989). Analysis of water distribution networks with head-dependent outlets. Civil Engineering Systems, 6 (3), 102-110

Rezaei, H., B. Ryan and I. Stoianov (2015). Pipe failure analysis and impact of dynamic hydraulic conditions in water supply networks. Procedia Engineering, 119, 253-262

Rezania, M., A. A. Javadi and O. Giustolisi, O. (2010). Evaluation of liquefaction potential based on CPT results using evolutionary polynomial regression. Computers and Geotechnics, 37 (1), 82-92

Rezania, M., A. A. Javadi and O. Giustolisi (2011). An evolutionary-based data mining technique for assessment of civil engineering systems. Engineering Computations, 25 (6), 500-517

Ripley, B. D. (2007). Pattern recognition and neural networks. Cambridge university press

Romano, M., Z. Kapelan and D. A. Savic (2014). Automated detection of pipe bursts and other events in water distribution systems. Journal of Water Resources Planning and Management, 140 (4), 457-467

Rogers, P. D. and N. S. Grigg (2009). Failure assessment modeling to prioritize water pipe renewal: two case studies. Journal of Infrastructure Systems, 15 (3), 162-171

Roshani, E. and Y. R. Filion (2013). Event-based approach to optimize the timing of water main rehabilitation with asset management strategies. Journal of Water Resources Planning and Management, 140 (6), 04014004

Rossman, L. A. (2000). EPANET 2 User's Manual. U.S. Environmental Protection Agency, Cincinnati, Ohio

Rossman, L. A. (2007). Discussion of "solution for water distribution systems under pressure-deficient conditions" by Wah Khim Ang and Paul W. Jowitt. Journal of Water Resources Planning and Management, 133 (6), 566-567

Røstum, J. (2000). Statistical modelling of pipe failures in water networks. PhD Thesis, Norwegian University of Science and Technology, Norway

Sadiq, R., B. Rajani and Y. Kleiner (2004a). Fuzzy-based method to evaluate soil corrosivity for prediction of water main deterioration. Journal of infrastructure systems, 10 (4), 149-156

Sadiq, R., B. Rajani and Y. Kleiner (2004b). Probabilistic risk analysis of corrosion associated failures in cast iron water mains. Reliability Engineering and System Safety, 86 (1), 1-10

Sadiq, R., B. Rajani and Y. Kleiner (2004c). Aggregative risk analysis for water quality failure in distribution networks. Journal of Water Supply: Research and Technology-Aqua, 53 (4), 241-261

Sadiq, R., Y. Kleiner and B. Rajani (2006). Estimating risk of contaminant intrusion in water distribution networks using Dempster–Shafer theory of evidence. Civil Engineering and Environmental Systems, 23 (3), 129-141

Sægrov, S., J. M. Baptista, P. Conroy, R. K. Herz, P. LeGauffre, G. Moss and M. Schiatti (1999). Rehabilitation of water networks: survey of research needs and on-going efforts. Urban Water, 1 (1), 15-22

Salehi, S., M. Jalili Ghazizadeh and M. Tabesh (2018). A comprehensive criteria-based multi-attribute decision-making model for rehabilitation of water distribution systems. Structure and Infrastructure Engineering, 14 (6), 743-765

Sander, A., B. Berghult, A. E. Broo, E. L. Johansson and T. A. Hedberg (1996). Iron corrosion in drinking water distribution systems—The effect of pH, calcium and hydrogen carbonate. Corrosion science, 38 (3), 443-455

Santos, P., C. Amado, S. T. Coelho and J. P. Leitão (2017). Stochastic data mining tools for pipe blockage failure prediction. Urban Water Journal, 14 (4), 343-353

Savic, D. A., O. Giustolisi and D. Laucelli (2009). Asset deterioration analysis using multi-utility data and multi-objective data mining. Journal of Hydroinformatics, 11 (3-4), 211-224

Sayyed, M. A. H. A., R. Gupta and T. T. Tanyimboh (2014). Modelling pressure deficient water distribution networks in EPANET. Procedia Engineering, 89, 626-631

Sayyed, M. A. H. A., R. Gupta and T. T. Tanyimboh (2015). Noniterative application of EPANET for pressure dependent modelling of water distribution systems. Water resources management, 29 (9), 3227-3242

Scheidegger, A., L. Scholten, M. Maurer and P Reichert (2013). Extension of pipe failure models to consider the absence of data from replaced pipes. Water research, 47 (11), 3696-3705

Scheidegger, A., J. P. Leitão and L. Scholten (2015). Statistical failure models for water distribution pipes–A review from a unified perspective. Water research, 83, 237-247

Scholten, L., A. Scheidegger, P. Reichert, M. Mauer and J. Lienert (2014). Strategic rehabilitation planning of piped water networks using multi-criteria decision analysis. Water research, 49, 124-143

Schuster, C. J. and E. A. McBean (2008). Impacts of cathodic protection on pipe break probabilities: a Toronto case study. Canadian Journal of Civil Engineering, 35 (2), 210-216

Selvakumar, A. and A. N. Tafuri (2012). Rehabilitation of aging water infrastructure systems: key challenges and issues. Journal of Infrastructure Systems, 18 (3), 202-209

Seni, G. and J. F. Elder (2010). Ensemble methods in data mining: improving accuracy through combining predictions. Synthesis Lectures on Data Mining and Knowledge Discovery, 2 (1), 1-126

Shahin, M. A., H. R. Maier and M. B. Jaksa (2004). Data division for developing neural networks applied to geotechnical engineering. Journal of Computing in Civil Engineering, 18 (2), 105-114

Shahnazari, H., M. A. Shahin and M. A. Tutunchian (2014). Evolutionary-based approaches for settlement prediction of shallow foundations on cohesionless soils. International Journal of Civil Engineering, 12 (1), 55-64

Shamir, U. and C. D. Howard (1979). An analytic approach to scheduling pipe replacement. Journal of the American Water Works Association, 71 (5), 248-258

Sheng, W. and X. Liu (2004). A hybrid algorithm for k-medoid clustering of large data sets. *In Proceedings of Evolutionary Computation Congress on IEEE (CEC 2004)*, 1, 77-82

Shin, H., K. Kobayashi, J. Koo and M. Do (2016). Estimating burst probability of water pipelines with a competing hazard model. Journal of Hydroinformatics, 18 (1), 126-135

Shirzad, A., M. Tabesh, R. Farmani and M. Mohammadi (2012). Pressure-discharge relations with application to head-driven simulation of water distribution networks. Journal of Water Resources Planning and Management, 139 (6), 660-670

Shirzad, A., M. Tabesh and R. Farmani (2014). A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks. KSCE Journal of Civil Engineering, 18 (4), 941-948

Shuang, Q., Y. Liu, Y. Tang, J. Liu and K. Shuang (2017). System Reliability Evaluation in Water Distribution Networks with the Impact of Valves Experiencing Cascading Failures. Water, 9 (6), 413

Siew, C. and T. T. Tanyimboh (2012). Pressure-dependent EPANET extension. Water Resources Management, 26 (6), 1477-1498

Singh, V. P., and J. Oh (2015). A Tsallis entropy-based redundancy measure for water distribution networks. Physica A: Statistical Mechanics and its Applications, 421, 360-376

Sitzenfrei, R., M. Mair, M. Möderl and W. Rauch (2011). Cascade vulnerability for risk analysis of water infrastructure. Water Science and Technology, 64 (9), 1885-1891

Sivakumar, P. and R. K. Prasad (2014). Simulation of water distribution network under pressure-deficient condition. Water resources management, 28 (10), 3271-3290

Skipworth, P. J., M. Engelhardt, A. Cashman, D. A. Savic, A. J. Saul and G. A. Walters (2002). Whole life costing for water distribution network management. Thomas Telford Publishing London

Smith, W. H. (1976). Frost Loadings on Underground Pipe. Journal of American Water Works Association, 68 (12), 673-674

Spickelmire, B. (2002). Corrosion considerations for ductile iron pipe. Materials Performance, 41 (7), 16-23

Stefanidis, S. and D. Stathis (2013). Assessment of flood hazard based on natural and anthropogenic factors using analytic hierarchy process (AHP). Natural hazards, 68 (2), 569-585

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. Journal of the royal statistical society. Series B (Methodological), 111-147

Tabesh, M., T. T. Tanyimboh and R. Burrows (2002). Head-driven simulation of water supply networks. International Journal of Engineering, 15 (1), 11-22

Tabesh, M., J. Soltani, R. Farmani and D. A. Savic (2009). Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling. Journal of Hydroinformatics, 11 (1), 1-17

Tabesh, M., and H. Saber (2012). A prioritization model for rehabilitation of water distribution networks using GIS. Water resources management, 26 (1), 225-241

Tang, Z. and B. McCabe (2007). Developing complete conditional probability tables from fractional data for Bayesian belief networks. Journal of Computing in Civil Engineering, 21 (4), 265-276

Tanyimboh, T. T., M. Tabeshand R. Burrows. (2001). Appraisal of source head methods for calculating reliability of water distribution networks. Journal of water resources planning and management, 127 (4), 206-213

Tanyimboh, T. T., B. Tahar and A. B. Templeman (2003). Pressure-driven modelling of water distribution systems. Water Science and Technology: Water Supply, 3 (1-2), 255-261

Tanyimboh, T. T. and A. B. Templeman (2010). Seamless pressure-deficient water distribution system model. *In Proceedings of the Institution of Civil Engineers (ICE)-Water Management*, 163 (8), 389-396

Tanyimboh, T. T., C. Siew, S. Saleh and A. Czajkowska (2016). Comparison of surrogate measures for the reliability and redundancy of water distribution systems. Water Resources Management, 30 (10), 3535-3552

Tesfamariam, S., B. Rajani and R. Sadiq (2006). Possibilistic approach for consideration of uncertainties to estimate structural capacity of ageing cast iron water mains. Canadian Journal of Civil Engineering, 33 (8), 1050-1064

Todini, E. (2000). Looped water distribution networks design using a resilience index based heuristic approach. Urban water, 2 (2), 115-122

Todini, E. (2003). A more realistic approach to the “extended period simulation” of water distribution networks, Advances in Water Supply Management, Balkema, Lisse, The Netherlands

Toumbou, B., J. P. Villeneuve, G. Beardsell and S. Duchesne (2012). General model for water-distribution pipe breaks: Development, methodology, and application to a small city in Quebec, Canada. Journal of Pipeline Systems Engineering and Practice, 5 (1)

Tscheikner-Gratl, F., R. Sitzenfrei, W. Rauch and M. Kleidorfer (2016). Enhancement of limited water supply network data for deterioration modelling and determination of rehabilitation rate. Structure and Infrastructure Engineering, 12(3), 366-380.

Tscheikner-Gratl, F., P. Egger, W. Rauch and M. Kleidorfer (2017). Comparison of multi-criteria decision support methods for integrated rehabilitation prioritization. Water, 9 (2), 68

Vairavamoorthy, K., J. Yan, H. M. Galgale and S. D. Gorantiwar (2007). IRA-WDS: A GIS-based risk analysis tool for water distribution systems. Environmental Modelling & Software, 22 (7), 951-965

Vamvakeridou-Lyroudia, L. S., J. Bicik, M. Morley, D. A. Savic and Z. Kapelan (2010). A real-time intervention management model for reducing impacts due to pipe isolation in water distribution systems. *In Proceedings of the 13th Water Distribution Systems Analysis Conference (WDSA) (pp. 209-221)*

Vanier, D. D. (2001). Why industry needs asset management tools. Journal of computing in civil engineering, 15 (1), 35-43

Vanrenterghem-Raven, A. (2007). Risk factors of structural degradation of an urban water distribution system. Journal of infrastructure systems, 13 (1), 55-64

Vreeburg, J. H. G., I. N. Vloerbergh, P. Van Thienen and R. De Bont (2013). Shared failure data for strategic asset management. Water Science and Technology: Water Supply, 13 (4), 1154-1160

Wagner, J. M., U. Shamir and D. H. Marks (1988). Water distribution reliability: simulation methods. Journal of Water Resources Planning and Management, 114 (3), 276-294

Walski, T. M. and A. Pelliccia (1982). Economic analysis of water main breaks. Journal of American Water Works Association, 74 (3), 140–147

Walski, T., D. Blakley, M. Evans and B. Whitman (2017). Verifying Pressure Dependent Demand Modeling. Procedia Engineering, 186, 364-371

Wang, Y., T. Zayed and O. Moselhi (2009a). Prediction models for annual break rates of water mains. Journal of performance of constructed facilities, 23 (1), 47-54

Wang, Y., O. Moselhi and T. Zayed (2009b). Study of the suitability of existing deterioration models for water mains. Journal of Performance of Constructed Facilities, 23 (1), 40-46

Wang, H. and X. Chen (2015). Optimization of maintenance planning for water distribution networks under random failures. Journal of Water Resources Planning and Management, 142 (2), 04015063

Ward, B., A. Selby, S. Gee and D. A. Savic (2017). Deterioration modelling of small-diameter water pipes under limited data availability. Urban Water Journal, 14 (7), 743-749

Watson, T., C. Christian, A. Mason, M. Smith and R. Meyer (2004). Bayesian-based pipe failure model. Journal of Hydroinformatics, 6 (4), 259-264

Wettschereck, D., D. W. Aha and T. Mohri (1997). A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. Artificial Intelligence Review, 11 (1-5), 273-314

Wilson, D., Y. Filion and I. Moore (2017). State-of-the-art review of water pipe failure prediction models and applicability to large-diameter mains. Urban Water Journal, 14 (2), 173-184

Winkler, D., M. Haltmeier, M. Kleidorfer, W. Rauch and F. Tscheikner-Gratl (2018). Pipe failure modelling for water distribution networks using boosted decision trees. Structure and Infrastructure Engineering, 1-10

Wood, A. and B. J. Lence (2009). Using water main break data to improve asset management for small and medium utilities: District of Maple Ridge, BC. Journal of Infrastructure Systems, 15 (2), 111-119

Wols, B. A. and P. van Thienen (2013). Impact of weather conditions on pipe failure: a statistical analysis. Journal of Water Supply: Research and Technology, 63 (3), 212-223

Wols, B. A. and P. van Thienen (2014). Modelling the effect of climate change induced soil settling on drinking water distribution pipes. Computers and Geotechnics, 55, 240-247

Wols, B. A., K. Van Daal and P. Van Thienen (2014). Effects of Climate Change on Drinking Water Distribution Network Integrity: Predicting Pipe Failure Resulting from Differential Soil Settlement. Procedia Engineering, 70, 1726-1734

Wu, Z.Y. (2007). Discussion of “Solution for Water Distribution Systems under

Pressure-Deficient Conditions''. Journal of Water Resources Planning Management, 133 (6), 567-568

Wu, Z. Y., R. H. Wang, T. M. Walski, S. Y. Yang, D. Bowdler and C. C. Baggett (2009). Extended global-gradient algorithm for pressure-dependent water distribution analysis. Journal of water resources planning and management, 135 (1), 13-22

Xu, C. and I. C. Goulter (1999). Reliability-based optimal design of water distribution networks. Journal of Water Resources Planning and Management, 125 (6), 352-362

Xu, Q., Q. Chen and W. Li (2011a). Application of genetic programming to modeling pipe failures in water distribution systems. Journal of Hydroinformatics, 13 (3), 419-428

Xu, Q., Q. Chen, W. Li and J. Ma (2011b). Pipe break prediction based on evolutionary data-driven methods with brief recorded data. Reliability Engineering and System Safety, 96 (8), 942-948

Xu, Q., Q. Chen, J. Ma and K. Blanckaert (2013). Optimal pipe replacement strategy based on break rate prediction through genetic programming for water distribution network. Journal of Hydro-Environment Research, 7 (2), 134-140

Yamijala, S., S. D. Guikema and K. Brumbelow (2009). Statistical models for the analysis of water distribution system pipe break data. Journal of Reliability Engineering and System Safety, 94 (2), 282-293

Yamini, H. and B. J. Lence (2010). Probability of failure analysis due to internal corrosion in cast-iron pipes. Journal of Infrastructure Systems, 16 (1), 73-80

Yazdani, A., and P. Jeffrey (2011). Applying network theory to quantify the redundancy and structural robustness of water distribution systems. Journal of Water Resources Planning and Management, 138 (2), 153-161

Yoo, D. G., M. Y. Suh, J. H. Kim, H. Jun and G. Chung (2012). Subsystem-based pressure dependent demand analysis in water distribution systems using effective supply. KSCE Journal of Civil Engineering, 16 (3), 457-464

Young, P., S. Parkinson and M. Lees (1996). Simplicity out of complexity in environmental modelling: Occam's razor revisited. Journal of applied statistics, 23 (2-3), 165-210

Zamenian, H., F. L. Mannering, D. M. Abraham and T. Iseley (2016). Modeling the frequency of water main breaks in water distribution systems: Random-parameters negative-binomial approach. Journal of Infrastructure Systems, 23 (2), 04016035

Zarghami, S. A., I. Gunawan and F. Schultmann (2018). Integrating entropy theory and cospanning tree technique for redundancy analysis of water distribution networks. Reliability Engineering & System Safety, 176, 102-112

Zhang, G., B. E. Patuwo and M. Y. Hu (1998). Forecasting with artificial neural networks: The state of the art. International journal of forecasting, 14 (1), 35-62

Zhou, Q. Y., J. Shimada and A. Sato (2001). Three-dimensional spatial and temporal monitoring of soil water content using electrical resistivity tomography. Water Resources Research, 37 (2), 273-285

Zweig, Mark H. and G. Campbell (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical chemistry 39 (4), 561-577