Assessing pipe condition in water distribution networks: deterministic versus heuristic approaches

Duarte Múrias Gray

Thesis to obtain a Master of Science Degree in

Civil Engineering

Supervisor: Prof. Dídia Isabel Cameira Covas

Examination committee

Chairperson: Prof. José Alexandre De Brito Aleixo Bogas,

Supervisor: Prof. Dídia Isabel Cameira Covas,

Members of the committee: Prof. Nelson Jorge Gaudêncio Carriço,

Dr. Marta Sofia Ferreira Cabral

July 2021

Declaration

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the *Universidade de Lisboa*.

Abstract

The condition assessment of water distribution pipes is of the utmost importance for the prioritization of rehabilitation interventions. However, available methodologies for condition assessment are very complex to be applied by water utilities that have few human, technological and financial resources. The current thesis aims at the development and application of a four-step methodology for the prediction of the physical condition of water distribution pipes without the need for visual inspection. The methodology includes two complementary approaches (deterministic and heuristic) considering different physical, operational and environmental factors that influence water pipe deterioration. The methodology is applied to the water distribution network of *Quinta do Lago* in the Algarve. Two types of regression models are implemented in the deterministic approach. Obtained results for the two approaches are compared with the performance indicator ratio of useful life, showing that the linear models in the deterministic approach tend to overestimate the pipe deterioration and the heuristic approach classifies the pipes mainly in average condition.

Keywords: condition assessment, water distribution pipes, deterministic approach, heuristic approach.

Resumo

A avaliação da condição de condutas de abastecimento de água é essencial para a priorização das ações de reabilitação necessárias. Contudo, as metodologias disponíveis para o efeito tendem a ser demasiado complexas para serem utlizadas de uma forma expedita por entidades gestoras, uma vez que as mesmas nem sempre possuem os meios humanos, tecnológicos ou financeiros para efetuar estas avaliações. A presente dissertação tem como objetivo o desenvolvimento de uma metodologia de quarto passos para a avaliação da condição de condutas de abastecimento e água sem a necessidade de inspeção ou interrupção do abastecimento. A metodologia desenvolvida possui duas abordagens complementares (determinística e heurística) que consideram diferentes fatores físicos, operacionais e ambientais para a contabilização da influência que os mesmos têm na deterioração. A metodologia apresentada é aplicada à rede de distribuição de água da Quinta do Lago no Algarve. Dois tipos de modelos de regressão são aplicados na abordagem determinística. Os resultados obtidos para as duas abordagens são comparados com o indicador do rácio de vida útil, que demonstra que os modelos determinísticos tendem a sobrestimar a deterioração das condutas e o modelo heurístico tende a classificar as condutas com uma condição média.

Palavras-chave: avaliação da condição, redes de distribuição de água, abordagem determinística, abordagem heurística.

Acknowledgments

For the development of this thesis I would like to thank my supervisor Professor Dídia Covas, for her guidance, perseverance throughout this entire process. I would like to thank her for her time and dedication and for constantly pushing forth the scientific content of this work.

I would like to thank FCT and the project WISDOM for the grant provided and for the feedback sessions that helped prepare the content of this investigation.

I would like to thank my family, especially my Mother and Father for their unconditional support of my studies and always giving me the motivation and drive I needed.

Finally, I would like to thank all my close friends for their companionship and friendship, and especially Andreia Freitas and Sofia Batista for their unwavering support to complete this degree.

Table of Contents

List of Figures

List of Tables

Abbreviations

AC	Asbestos concrete
AMP	Asset management plans
ANN	Artificial neural network
CAD	Computer assisted design
CCTV	Closed-circuit television
DI	Ductile iron
DN	Nominal pipe diameter
GIS	Geographical information system
GPR	Ground-penetrating radar
HDPE	High density polyethylene
IVI	Infrastructure value index
MFL	Magnetic flux leakage
PVC	Polyvinyl chloride
RFEC	Remote field eddy current
RUL	Ratio of useful life
SONAR	Sound navigation radar
UN	United Nations
UWB	Ultra-wideband
VIF	Variance inflation factor
WRc	Water research centre

Notation

Symbol	Description	Units
C_n	Condition value for the pipe *n*	[-]
$CS_{i,t}$	Cost of replacement of the asset *i* at the time *t*	€
cov(x,y)	Covariance of the variables *x* and *y*	[-]
f_i	Fitted value predicted	[-]
IVI	Infrastructure value index	[-]
IDI	Infrastructure deterioration index	[-]
N	Number of data points	[-]
$N_{n,j}$	Normalized and weighted influencing factor *i* for pipe *n*	[-]
PO	Period of observation (years)	years
P	Number of dependents (or explanatory) variables	[-]
R^2_{adj}	Adjusted r-squared	[-]
R^2	r-squared	[-]
r_s	Spearman rank coefficient	[-]
$rul_{i,t}$	Residual life of the pipe *i* in the year *t*	[-]
t	Time	years
VIF_i	Variance inflation factor for the variable *i*	[-]
$vr_{i,t}$	Remaining service life (years) of the asset *i* at the time *t*	years
vu_i	Service life of the asset $_i$	years
W_i	Attributed weight for the influencing factor	[-]
$\overline{y}$	Mean of the data values y	[-]
$y_{n,i}$	Value of the influencing factor *i* for pipe *n*	[-]
$y_{min,i}$	Minimal value for the influencing factor *i*	[-]
$y_{max,i}$	Maximum value for the influencing factor *i*	[-]
σ_x	Standard deviation for the variable *x*	[-]
σ_Y	Standard deviation for the variable $_Y$	[-]
$\rho_{X,Y}$	Pearson's product-moment-correlation-coefficient between the variables *x* and *y*	[-]

1 Introduction

1.1 Context and motivation

The United Nations (UN) established, in 2015, a set of 17 global development goals that aim to provide peace, dignity and prosperity to all nations of the world. These goals were established by world leaders as a blueprint to improve the world we live in and to prepare for future challenges. The 11th goal is to "*make cities inclusive, safe, resilient and sustainable*" and one of the items in the agenda to achieve this goal is "*improving human settlement management*", as well as "*promoting the integrated provision of environmental infrastructure: water, sanitation, drainage and solid waste management"* (United Nations, 2015). The consequences that these goals have on policy making and society are fundamental for cooperative behaviour and the resolution of global issues. Additionally, these goals not only propose a clear course of action for governments and organizations, but also validate all underlying motivations behind these goals.

The 11th goal requires the understanding of the definition of sustainability, which is defined in ISO 82: 2019 as the *"state of the global system, including environmental, social and economic aspects, in which the needs of the present are met without compromising the ability of future generations to meet their own needs".* In the field of urban water systems, sustainability, along with sustainable development, is dependent on the type of economy. While emerging economies are focused on building new infrastructures to supply drinking water and collect wastewater to population and to cope with rapid growth, developed economies are focused on maintenance and rehabilitation of existing infrastructures (Clark *et al.*, 2002). In developed economies, the current challenges in the urban water field are the management of ageing infrastructures (Cabral *et al.*, 2019) and the effects it has on the service provided (Alegre & Covas, 2010). Consequently, infrastructure asset management has become one of the priorities of water utilities, with special focus on the rehabilitation of pipe networks and facilities and on the efficient management of funds by optimizing resource allocation through efficient prioritization of infrastructure improvements (Najafi & Perez, 2016).

Condition assessment of existing assets is of the utmost importance to correctly evaluate the current state of infrastructures and their assets and to aid in the implementation of infrastructure asset management, which include investment and Operation & Maintenance (O&M) plans (Al-Barqawi & Zayed, 2006a; Alegre & Covas, 2010). Current methods to evaluate the physical condition of buried infrastructures can either require direct inspection or the interpretation of operational, environmental and physical factors to indirectly predict pipe condition. The latter approach is also known as indirect condition assessment (Al-Barqawi & Zayed, 2006b). Direct inspection methods are often used for large diameter pipes, as they are not adequate for pipes with a nominal diameter less than 250 mm, due to operational constraints and also the small economic cost of failure of small-diameter pipes in comparison to large diameter pipes (Thomson & Royer, 2013).

Nevertheless, condition assessment of small diameter pipes is still very important and cannot be neglected to successfully prioritize investment and rehabilitation works in water distribution networks. Several indirect condition assessment methods have been proposed, being, most methods highly reliant on the construction of advanced computer aided systems, such as machine learning algorithms, that have shown increasingly accurate results (Al-Barqawi & Zayed, 2006a; Berardi *et al*., 2008; Tuhovčák *et al*., 2013). However, these methods are very challenging to implement in water distribution systems, given the limited infrastructure data (St. Clair & Sinha, 2012). In Portugal, 51 out of 136 water utilities do not have any form of formal inventory of assets, including information related to pipe location and historical data, and only 20 water utilities reported to have any information regarding the condition of the assets (RASARP, 2020).

Overall, there is an urgent need to address the gap between the existing condition assessment methodologies and the existing required resources in water utilities, namely, financial, human and technological. Regardless of the research advances in the field, the need to develop a sound methodology to indirectly assess pipe condition, yet easy to apply by water utilities, is the main motivation of this research.

1.2 Objectives and methodology

The main objective of this research is to develop a methodology to assess the physical condition of water distribution networks, without the need for visual inspection or service interruption. The proposed methodology takes into consideration the data gaps that Portuguese water utilities face, regarding the scarce infrastructure knowledge and the difficulty of using the current indirect condition assessment methods. Two approaches (deterministic and heuristic) are developed, compared and applied to assess pipe condition of Quinta do Lago water distribution network.

This research is a part of the project WISDOM (Water Intelligent System Data) that aims to develop algorithms and models that use data collected by water utilities to draw conclusions that help in decision making.

This thesis is organised in the following steps:

i. Literature review of the existing knowledge on pipe condition assessment and the tools already used in the urban water field.
ii. Development of a methodology to assess the physical condition of water distribution networks, using deterministic and heuristic approaches.
iii. Application and validation of the proposed methodology to the water distribution system of Quinta do Lago.
iv. Comparison of the obtained results by the two approaches, final considerations and suggestions for future research.

1.4 Outline

This thesis is organized into eight chapters. The first presents an introduction to the study, in which the relevance and motivation for this investigation are explained, as well as the objectives and the methodology chosen. In addition, this chapter also contains an outline of the structure of this dissertation.

Chapter 2 presents the literature review on the main issues related to condition assessment. Approaching the general aspects of condition assessment as well as defining the relevant terms associated. This chapter also describes the most common direct and indirect condition assessment methods, highlighting the relevant aspects of each.

Chapter 3 describes, in detail, the proposed methodology for pipe condition assessment by using the deterministic and heuristic modelling approaches.

Chapter 4 includes the description of the case study analysed in this research. The chapter also includes the discussion of relationships between the network properties (i.e., year of installation and pipe material) and the validation of the collected data.

Chapter 5 presents the application of the deterministic approach to the case study, detailing the attempts made to model physical condition.

Chapter 6 describes the data collection and results for the heuristic modelling approach as well as the application of this approach to the network of Quinta do Lago.

Chapter 7 compares the results of the heuristic and deterministic modelling approaches, highlighting the similarities and differences between them.

Chapter 8 includes the final remarks of this investigation, as well as the conclusions about the different approaches used and the future works suggested.

2 Literature review

2.1 Introduction

The current chapter aims to present the state-of-the-art review about methodologies for condition assessment of water pipes in order to identify the most cost-effective methods and to discuss their advantages and disadvantages. Firstly, this chapter presents the scope of condition assessment, discussing its definition, importance in infrastructure asset management and the frameworks used, as well as the importance of an appropriate condition rating scale. A description of the most common condition assessment methods is presented, divided in two sections (i.e., direct and indirect methods), providing a discussion on which scenarios each method is most appropriate. Finally, the chapter addresses the current knowledge gaps which this investigation aims to bridge.

2.2 Condition assessment

When assessing the *condition* of a water distribution pipe, the term *condition* can refer to the hydraulic capacity, the economic evaluation, the structural integrity or a combination of any of these factors (Thomson & Royer, 2013). The hydraulic capacity and the economic evaluation condition require the review of external variables, such as the population needs, water pressure logs, water velocity logs and estimation of replacement costs (Cabral *et al.*, 2019; Vidigal *et al.*, 2010). On the other hand, condition assessment of the pipe structural integrity refers exclusively to its properties (i.e., pipe age, material, nominal diameter) and the external factors (i.e., soil pH, surface traffic, pressure variations) (Howard, 2001). In this study, the term *condition* will only refer to the structural integrity of the evaluated pipes, also referred to as the pipe physical condition. This is defined *"as the presence or absence of holes, cracks or the conditions leading up to their formation, in transmission of distribution pipe wall, lining, coating and joints*"(Field *et al.*, 2007).

The assessment of the physical condition of assets is an important part in infrastructure asset management (IAM), which corresponds to a set of processes that can be applied by water utilities to achieve adequate infrastructure cost, risk and performance (Alegre & Coelho, 2012). This is aligned with the definition of asset management in ISO 55000: 2014. Furthermore, IAM is the broader scope for decision making, aiming to reduce the life-cycle costs of assets, which is defined as all costs related to the asset during its lifetime with exception of the planning and design stages (ISO 15686-8: 2014). An efficient IAM aims to provide an asset management plan (AMP), allowing to minimize the life-cycle costs by: lowering renewal costs, identify the necessary measures for optimized resource allocation and prioritize the necessary interventions (Harvey *et al.*, 2017). Thus, an comprehensive methodology for IAM that considers decisions on strategic, tactical and operational level for asset management is needed (Alegre & Covas, 2010).

Condition assessment is integrated into IAM in the process of asset performance evaluation. Often, key performance indicators (KPI) are used by water utilities to evaluate performance and to decide about the improvement actions (Bank, 2013; Covas *et al.*, 2018; Royer, 2012). However, condition assessment allows to estimate the remaining service life, given the associated risks of failure and the future needs of the population, as well as work alongside risk management, maintenance planning and data collection techniques to achieve robust and accurate results (IPWEA, 2006).

The process of condition assessment is, generally, known to water utilities and regulation authorities, however, the used methodologies experience a variety of forms. A general framework to evaluate the physical condition of a water distribution system established by United States Environmental Protection Agency (EPA) is described in Figure 2-1. This framework explains the three key aspects of condition assessment: condition rating scale, data collection and analysis. Although the author, Field *et al.* (2007), presents the aspects in an order only suitable for the specific study, illustrates that condition assessment requires a condition rating scale to describe the results of the assessment method. Furthermore, the importance of data collection and pipeline characterization allows for an adequate investigation of the main factors that might be affecting the deterioration rates of the water distribution pipes. The holistic understanding of the distribution network is very important in decision making processes, which can only be achieved by the collection and analysis of the influencing factors of deterioration (e.g., burst rate, pipe material, operating pressures). Finally, the analysis of the asset via a direct or indirect method provides necessary tools to achieve the desired results of condition assessment (Field *et al.*, 2007; Marlow & Burn, 2008).

1) Development of a structural-condition-rating scale

Development of a formal or informal structural-condition-rating that links pipe parameter data to the likelihood of structural failure (i.e., holes, cracks, and breaks) for the time period of interest.

2) Collection of data and pipeline characterization

Collection of data related to physical, envirmomental and operational factors effecting the asset, along with historical data of failure history, operational processes and associated indicators, acquired via direct or indirect methods of assessment for pipeline characterization

3) Analysis

Analysis of the pipeline data and information aquired, into categorize pipe structural condition, based on the structural condition rating previously defined. As well as, the likelihood of failure for the pipeline sections and time periods of interest. The structural condition assessment results are used as input to an informal or formal decision-support system.

Figure 2-1 – Framework for condition assessment proposed by Field, *et al.*(2007)

The output result of any condition assessment method is generally given in a condition rating scale or condition index, which allows to quantify the asset condition (Amani *et al.*, 2012; IPWEA, 2016) and to identify priority assets. The adopted condition rating scale is usually carried out by using, or adapting, a known rating scale established by a regulation authority, a research institution or a utility with recognised best-practices. Condition rating scales are essential in any framework of performance evaluation by virtue of *"their ability to form a basis for measuring deterioration and prediction of each component or facilities"* (Amani *et al.*, 2012). Furthermore, the condition rating scale can be in the form of a discrete scale, for example from 1 to 5, however, in some cases a continuous scale can also be used. The chosen type of scale is mostly dependent on the output provided by the condition assessment method. Meaning that, if the method chosen produces results in the form of a continuous scale, the rating scale can also be defined in a continuous form. However, the condition rating scales presented below go beyond the categorization of the physical condition of the asset (i.e., very good, good, inadequate), and associate the physical condition with necessary maintenance action or the predicted maintenance scheduling required.

Two discrete condition rating scales are presented in Table 2-1 and in Table 2-2, both proposed by national institutions to help their local municipalities in their condition assessment effort. Different descriptions for the same levels of condition are presented which shows that each institution prioritizes different results. While both rating scales propose a form of qualitative assessment, either by, number of defects of the overall state of asset. The grading scale proposed by U.S. Environmental Protection Agency, (2015) presents a timeline for failure, while the scale proposed by IPWEA, (2016) presents a level of required maintenance. Both scales have advantages and drawbacks, although a failure timeline can be very useful, it assumes that the maintenance practices of every water utility is the same, which is not the case. On the other hand, a continuous grading scale proposed by Al-Barqawi and Zayed (2006b) is presented in Table 2-3. The condition rating scale is established to be appropriate for the specific results of the condition assessment method proposed by the author. The relevance of this scale is it demonstrates that, although the development of a grading scale can be associated with regulatory institutions, it must be adapted to the assessment method.

Table 2-1 – Severity grading scale for condition assessment proposed by U.S Environmental Protection Agency (2015).

Severity grade	Description
1	Pipe segment has minor defects - failure unlikely in the foreseeable future
2	Pipe segment has minor defects - pipe unlikely to fail for at least 20 years
3	Pipe segment has moderate defects - deterioration may continue, at a ten to a twenty-year timeframe
4	Pipe segment has severe defects - risk of failure within the next five to ten years.
5	Pipe segment has failed or will likely fail within the next five years - requires immediate attention

Table 2-2 – Grading scale for condition assessment proposed by IPWEA (2006).

Rank	Description
1	Very good condition – only regular maintenance required
2	Minor defects only – minor maintenance required (5% needs replacement)
3	Maintenance required to return to accepted level of service significant maintenance required (10-20% needs replacement)
4	Requires renewal significant renewal/upgrade required (20-40% needs replacement)
5	Asset unserviceable (Over 50% of asset requires replacement)

Table 2-3 – Continuous grading scale proposed by Al-Barqawi and Zayed (2006b) for condition rating of water main.

Cardinal scale	Nominal scale	Criteria	Action
0-3	Critical	Severe internal or external corrosion. The remaining wall thickness is less than 50% of the original.	Immediate repair or replacement required.
3-4	Poor	No lining or coatings. Significant signs of internal or external corrosion. Remaining wall thickness 50 to 75% of the original.	Schedule for rehabilitation or replacement within the next 3–5 years.
4-6	Moderate	Some damage to coatings and/or linings was observed. Remaining wall thickness 75% or more of the original.	Reassess in 3–5 years. Schedule for lining and rehabilitation within the next 5–10 years.
6-8	Good	Coatings, linings still intact. Remaining wall thickness more than 90% of original	Reassess in 10 years Schedule for cathodic protection within the next 5–10 years.
8-9	Very good	Like new with no signs of corrosion or deterioration	Reassess in 15 years.
9-10	Excellent	Newly/recently installed.	No action is required.

2.3 Overview of condition assessment methods

The condition assessment of assets in a water distribution can be only be performed, in most cases, using off-line methods (i.e., methods that are not fed by continuous information sources), namely direct or indirect condition assessment methods, which are a reliable option for this type of assessment (Urrea-Mallebrera *et al.*, 2019). To evaluate the physical condition of any urban asset (i.e., buildings, treatment plants, roads), the use of visual inspection as a method for condition assessment is very common as a form of preliminary inspection. The visual inspection of an ass*et al*lows for symptomatic analysis and the ranking of the asset deterioration, which, in some cases, can be quantified using *in-situ* measurements (Flores-Colen *et al.*, 2008). However, the inspection of buried infrastructures is more complex, given that it may require specialized equipment and possibly the interruption of service (Royer, 2012).

Nevertheless, to understand the scope of condition assessment, the key differences between methods and methodologies should be analysed. Firstly, methodologies are the overall scope of an approach and the assumptions made to achieve the desired output, whereas methods are the tools used to achieve outputs (Azevedo *et al.*, 2011; Reddy, 2018). Condition assessment methodologies have advanced from reactive (or corrective) programs to proactive (or preventive) programs, providing direct benefits to water utilities (e.g., reduction in O&M costs, water loss reduction), as well as indirect benefits to the consumers and the society in general (e.g., reduction of traffic congestion, reduction in economic losses), and improving infrastructure sustainability (e.g., reduction in property damage, reduction of public distrust costs) (Jin Jun *et al.*, 2020).

Moreover, the methods used in condition assessment are the tools used to acquire information related to distribution systems and can be divided into two main groups (Figure 2-2): direct and indirect condition assessment methods. While direct condition assessment methods require the inspection of the asset, indirect condition methods collect and analyse the assets' physical characteristics (e.g., pipe material, pipe diameter, pipe age), failure history (e.g., number of busts per year, burst location) and the previous inspection information, in case this is available, to predict the physical condition (Liu & Kleiner, 2012) (Al-Barqawi & Zayed, 2006b). Additionally, direct condition assessment methods can be intrusive (i.e., requiring access to the interior of the pipe with or without interruption of service) and non-intrusive methods (i.e., not requiring access to the interior of the pipe, but often requiring excavation) (Royer, 2012). On the other hand, indirect condition assessment methods are categorized in environmental (i.e., involving the surveying of soil and water chemistry) and operational (i.e., involving leak/burst history, pipe material, embedment, coating, or linings) and aim to analysed trends and changes in the system to provide information regarding the condition of assets (Selvakumar *et al.*, 2013; Ugarelli & Bruaset, 2013). The use of the direct methods is not recommendable in non-critical pipes of water distribution networks, due to the high costs associated (Royer, 2012; Rajani and Kleiner, 2001b), nor in small

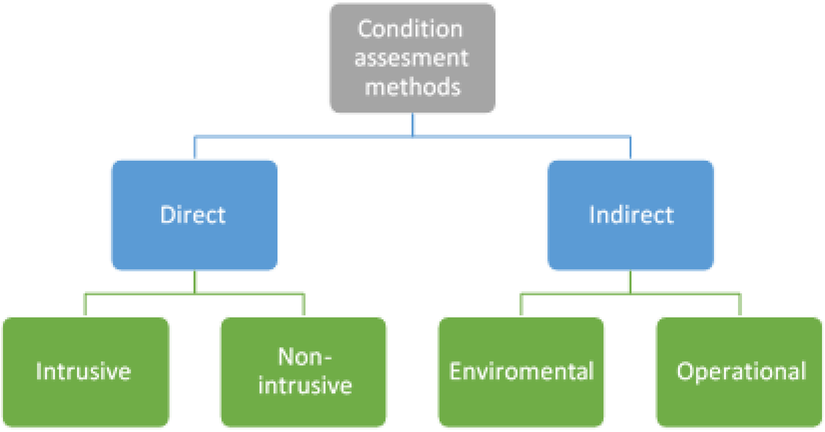


Figure 2-2 – Categories of condition assessment methods (Ugarelli & Bruaset, 2013).

diameter pipes since the equipment cannot be inserted inside the pipe.

2.4 Direct condition assessment

2.4.1 Introduction

The direct condition assessment methods require some form of inspection of the pipeline (Al-Barqawi & Zayed, 2006b). The underlying strategy used by water utilities to schedule inspection can be proactive or reactive (Rizzo, 2010). While proactive strategies aim to prevent pipe bursts, reactive strategies aim to minimize the reaction time taken to detect and repair the pipe burst and to re-establish the supply service in order mitigate the impacts of the supply disruption. From the economical point of view, proactive strategies have proven to be the most cost-effective in asset management plans, reducing the overall cost associated with pipe inspection and burst repair (Jin Jun *et al.*, 2020).

The following sections of the literature review will present with the most relevant types of direct condition assessment methods. These sections will not be an exhaustive list of all methods but will describe and explain the leading principles of the most used methods.

2.4.2 Closed-circuit television cameras

Closed-circuit television (CCTV) is a type of direct condition assessment method, classified as an intrusive method. For the implementation of this method, the pipe can be empty, partially empty or pressurized, depending on the specifications of the CCTV inspection technology. This method can be used in all types of pipe materials (Ruchti, 2017). The used equipment comprises a television camera set on a robot, which is assisted by an illumination system that is placed inside the pipe via a manhole. The robot, also called a crawler, moves along the pipe and is manually operated from the surface by an operator (Lampola & Kuikka, 2018).

This inspection method is usually very time-consuming, since the robot must stop at each detected defect in the pipe for rigorous evaluation by the operator, pinpointing its location and dimension. The time consumed per inspected section is directly proportional to the deterioration of the pipe. This inspection method is commonly used in wastewater pipes (Feeney *et al.*, 2009; Zangenehmadar, 2016).

In smaller diameter pipes, the commercially available robots that carry the inspection camera do not fit inside the pipe, as a result, the operator must use a *push-cam* method to backtrack from the pipe end to the point of entry. However, this method severely compromises the analysis, because it does not allow the operator to freely move the camera view and thoroughly inspect the pipe (Feeney *et al.*, 2009). The push-cam is also used in pressurized systems, in which a camera is pushed through a chlorinated seal, usually at the location of fire hydrants, and used to inspect the inside of the pipe removing the need to turn off the service provided (Figure 2-3).

The assessment of the degree of pipe deterioration significantly depends on the experience and judgment of the operator (Zangenehmadar, 2016). Many authors have criticized this method when aiming to evaluate an entire network, because it does not yield objective analyses and leads to imprecise results (Lampola & Kuikka, 2018; Zangenehmadar, 2016; Feeney *et al.*, 2009).

A working CCTV system inspecting a water distribution pipe through a fire hydrant entry is depicted Figure 2-3. Figure 2-3(a) presents the chlorinated seal and Figure 2-3(b) the view of the controller from the surface monitor.

CCTV video inspection is the most frequently used direct inspection method (U.S. Environmental Protection Agency, 2015), because it is the most cost-effective, providing a broader range of applications with the least associated costs. Additionally, CCTV inspection is the only methods that can provide visual data on defects, regarding location, sediment and debris build-up (Feeney *et al.*, 2009).

(a) (b)

Figure 2-3 – Working CCTV camera system for condition assessment: (a) fire hydrant seal; and (b) controller view (API -Advanced pipe inspection UK Ltd., 2021).

2.4.3 Laser scan

The laser scanner is classified as a direct intrusive condition assessment method, being more effective when combined with the CCTV method, because of its high accuracy quantitative results (Tavakoli, 2018). However, this method can only be used in large diameter pipes (> 250 mm), due to the size of the robot with the laser. The operational steps for laser scanning inspection methods are similar to those of the CCTV: the laser scanner is inserted inside the pipe through the entry point and guided along the pipe using laser light beams to measure the distance between the laser source and the pipe wall. The distance measured by the laser can be calculated using triangulation, time-of-flight, pulse-type time-of-flight or modulated beam systems (Royer, 2012).

The advantage of this method is that the output results can be used to create 2D or 3D maps of the inside the pipe, allowing for the evaluation of geometric discrepancies that might indicate a leak or a crack. This method is usually used in very long pipes and has an accuracy error in defect location of 0.5% (Lampola & Kuikka, 2018; Selvakumar *et al.*, 2013). However, the inspection of the pipe through laser scanning only provides a "point cloud" that needs to be processed and converted into a drawing format (e.g., CAD). Only, then, can the association between the CCTV results and the laser scan data be achieved (Figure 2-5) (Arrival 3D, 2020; Garvey, 2012). Finally, being the CCTV method already considered a very time-consuming method, the association of this equipment with a laser scanning tool to improve the results is unfeasible to apply in the entire water distribution network.

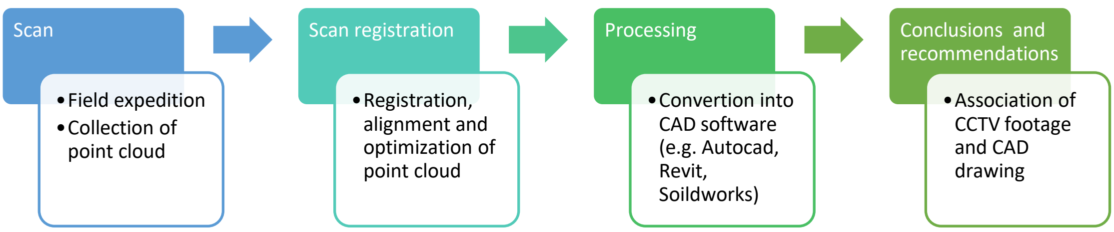


Figure 2-4 – Sets of condition assessment method using laser scanning and CCTV footage laser (Arrival 3D, 2020; Garvey, 2012).

2.4.4 Magnetic flux leakage

The magnetic flux leakage (MFL) condition assessment method is classified as a direct intrusive method in virtue of its equipment requiring be in direct contact with an empty pipeline. This method is generally used in large diameter metallic pipes, mostly in oil transmission pipelines (Shi *et al.*, 2015). The electromagnetic phenomenon behind this method is that a constant magnetic field is disturbed in the presence of defects. Consequently, when the constant magnetic field passes through a defect in the analysed pipe, the magnetic field changes due to different pipe properties (e.g., conductivity, thickness, shape) and the registered change can be recorded and interpreted as a defect. The constant magnetic field is described as an electro-magnet and moves along the pipe where these defects change the electromagnetic field (Tavakoli, 2018). This method is hardly applied in the urban water field, as it is limited to metallic clean unlined pipes. However, this condition assessment method allows to fully detect defects in the interior and exterior of the pipe (Liu & Kleiner, 2012; Shi *et al.*, 2015).

2.4.5 Remote field eddy current

The remote field eddy current (RFEC) method is classified as a direct intrusive condition assessment method. This method is similar to magnetic flux leakage method, as it is only limited to ferromagnetic materials. However, RFEC method can use smaller equipment, which makes it more appropriate to use in small diameter pipes and allows them to work in pressurized systems given the correct seal technology (Royer, 2012; Zangenehmadar, 2016). As presented in Figure 2-5, to perform this method, exciting and detection coils are placed inside the pipe at a distance of 2.5 to 3 diameters of the inspected pipe. The exciting coil transmits a low-frequency magnetic field which propagates outwards and reaches the detector coil either by a direct or indirect energy path. At the set distance, the signal received via the direct pathway is attenuated, while the signal traveling via the indirect transmission path is much stronger. Furthermore, the signal received from the indirect pathway passes through the pipe wall at the exciter and detector coil, allowing to measure the relationship between the signal of the two positions and to detect the percentage of missing material, leading to the conclusion regarding the pipe deterioration (Duchesne *et al.*, 2011).

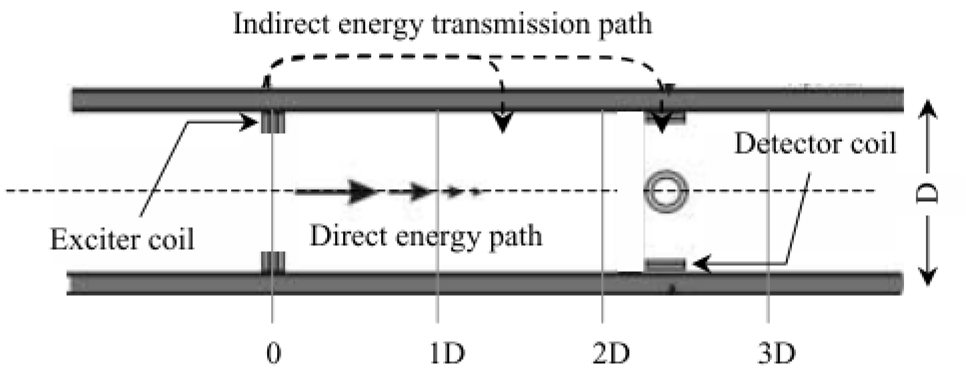


Figure 2-5 – Remote field eddy current in a metallic pipe (Rajani & Kleiner, 2004).

This method is reported to be very useful in analysing the condition of water supply networks, but requires a very skilled technician to interpret the amplitude and phase signals (Zangenehmadar, 2016). However, this method does not require the cleaning of the pipe to remove blockages caused by tuberculation or sedimentation. The advantage of this method, in addition to the decrease in the labour involved, is that it does not need the removal of tuberculation of the pipe walls and, consequently, does not damage pipes during the cleaning process (AWWA, 2014).

2.4.6 Broadband electromagnetic

Broadband electromagnetic (BEM) condition assessment method is classified as a direct intrusive method, and like the remote field eddy current method, the pipe wall thickness is calculated through the signal attenuation that travels via the indirect pathway between the transmitter and the detector (Mazumder *et al.*, 2018). However, the BEM covers more than a single frequency, unlike the remote field eddy current method. Likewise, the signal frequency ranges between 50 Hz and 50 kHz which allows the detection and quantification of several wall thicknesses as well as the effective conductivity of the pipe (Royer, 2012). Additionally, when a change in conductivity happens as a result of decreased signal detection, this indicates changes in material properties (i.e., wall thickness or leak) (Tai *et al.*, 1996).

However, BEM assessment requires the interruption of service and the cleaning of inner surface of the pipe wall in order not to interfere with data readings. The technology can examine pipe condition through the internal and external coating, but the method is considered to be slow and expensive (Mazumder *et al.*, 2018). However, some commercially available BEM methods can carry out the assessment intrusively without the need for interruption of the service. Recent technological advances in this process have allowed BEM equipment to become portable and access pipe exterior through small diameter perforation (Liu & Kleiner, 2012).

2.4.7 Ground-penetrating radar

Ground-penetrating radar (GPR) is one of the most popular near-surface geophysical methods used for infrastructure inspection (Wai-Lok Lai *et al.*, 2018). Originally developed by the U.S. Military to locate underground tunnels and landmines (Feeney *et al.*, 2009), GPR uses antennae to transmit electromagnetic pulses through the ground that reflect off objects and can be picked up by detectors and interpreted by a specialist (Liu & Kleiner, 2012). The range frequency of a typical GPR instrument is between 10-5000 MHz, which allows for GPR to be very versatile in its application and less influenced by electromagnetic noise. The popularity of GPR can be closely associated with the immediate on-site feedback that the technology can provide, as well as the ability to be processed and interpreted to create a 3D model of the subsoil infrastructure (Wai-Lok Lai *et al.*, 2018).

GPR can be classified as a direct non-intrusive condition assessment method, because it does not require access to the interior of the water distribution system. Its main application in pipe condition assessment is the ability to detect leaks and voids that have been created by water.

Furthermore, its 3D capabilities allow this method to be used in all materials and with a very high degree of accuracy in asset mapping (Zangenehmadar, 2016).

However, although its commercial availability is an advantage for the usage of this technology (Wai-Lok Lai *et al.*, 2018), this technology is not significantly used for pipe inspection, mainly because it is described to be at a prototype development stage (Feeney *et al.*, 2009) and results are highly dependent on surrounding soil conditions (i.e., density, water level) (Mazumder *et al.*, 2018; Royer, 2012).

2.4.8 Ultra-wideband pulsed radar system: P-Scan

Ultra-wideband pulsed radar system (UWB) can be considered a direct non-intrusive method of condition assessment. Mostly applied in non-metallic pipes for leak detection. The technology works by the transmission and detection of pico and nano-second waves that, like previous electromagnetic methods, use the feedback to determine location, orientation and geometry of the buried assets and voids in the pipe surround soil and, as a result, infer about the pipe overall condition (Zangenehmadar, 2016).

Although this is not a new technology, having already been successfully used since 1960, its market penetration is not extensive and commercially available options are not frequent. The reason behind this might be due to the regulations towards the emission of electromagnetic waves in certain wavelengths that can interfere with telecommunication signals (Allouche, 2017).

2.4.9 Sonar profiling system

Sonar (<u>so</u>und <u>na</u>vigation and <u>r</u>anging) profiling system is an acoustic method of condition assessment created to be used underwater (Feeney *et al.*, 2009). Similarly, to other intrusive methods, the equipment needs to be placed inside the pipe to operate, therefore, it is considered to be a direct intrusive method. The sonar device can either be an underwater scanner unit or a sonar siphon float depending if it is a water distribution or a wastewater pipe, respectively. The device uses the acoustic waves to register the cross-section of the pipe (Figure 2-6) (Liu & Kleiner, 2012). When inspecting partially empty water pipes, the speed of acoustic waves emitted from the sonar device travels at different speeds in water and air. Therefore, to analyse the pipe as a whole the two images have to be merged to create the full profile (Zangenehmadar, 2016).

The device travels through the pipe at a speed of 0.1 and 0.2 m/s and emits an acoustic wave every 1.5 s. Using high-frequency sonar allows for higher resolution pipe profiling, but has low penetration power and some inaccuracies are not registered. Low-frequency sonar devices have the inverse problem: low resolution with high penetration. As a result, the use of high-frequency sonar is most common in distribution systems, while in wastewater systems, low frequency sound waves are often a better alternative (Liu & Kleiner, 2012). The use of sonar and CCTV systems in pipe inspection is a good alternative to mitigate the errors associated with the subjectivity of the operator. However, it is not a perfect system, as sonar systems have difficulty in detecting longitudinal cracks and require an experienced technician to be operated (Feeney *et al.*, 2009; Mazumder *et al.*, 2018).

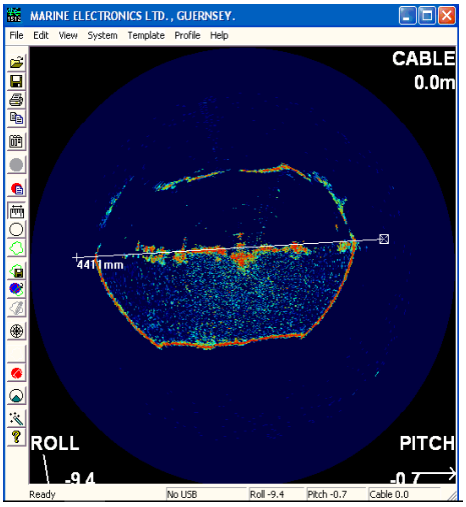


Figure 2-6 – Cross-section of pipe using SONAR

2.4.10 SmartBall®

The Smartball® is a relatively new system developed by PURE technologies in 2005 that incorporates a variety of sensors such as acoustic sensors, accelerometers, ultrasonic transmitters and temperature sensors inside an aluminium shell (Zangenehmadar, 2016). The outer layer of the ball is made of polyurethane foam, making the ball naturally buoyant and able to be inserted into full, or partially full, water pipes (Feeney *et al.*, 2009). The size of the ball is 1/3 the diameter of the pipe and travels through the entire pipe and transmits data every 3 seconds. This technology is a direct intrusive method (Zangenehmadar, 2016). The main advantage of this technology is that it can quickly examine an extensive pipeline by allowing the equipment to be deployed for 16 to 18 h at a time, obtaining high-quality results and no service interruption. However, the equipment is hard to control once inside a distribution system (Pure Tecnhologies, 2016). Figure 2-7 shows a Smartball being tested for quality control in a replica of an actual water system (Xylem, 2020).

Figure 2-7 – Smartball® developed by PURE technologies (Xylem, 2020).

2.4.11 Guided wave ultrasound

The guided wave ultrasound method belongs to a group of methods that use ultrasonic principles and equipment to carry out inspections. The physical principle behind this method is that, like sound waves, ultrasonic waves can travel along the pipe thickness to echo off defects in the pipe wall and upon returning to the transducer where the ultrasonic waves are converted into electromagnetic signals to be interpreted. Upon receiving the signal, the time-lag is recorded and given the propagation speed of the wave through the pipe, the distance to defect from the emitter can be determined (Liu & Kleiner, 2012). This is performed through the exterior of the pipe and only a small section of the pipe needs to be exposed to install the probe (Royer, 2012). Given this operational description, the method can be described as a non-intrusive direct condition assessment method.

However, dispersion curves need to be calculated for each type of wave before using this method. Dispersion curves describe the relationship between frequency, phase and thickness. To calculate these curves the governing equations are Navier-strokes, which coupled with the assumptions made about pipe boundary, can provide the harmonic solutions which are related to the phase velocity, frequency and, ultimately, to the wall thickness (Rose, 2003).

For the detection of longitudinal cracks, torsional waves are used to create shear stresses along the crack and obtain an echo back at the receiver (Lowe & Cawley, 2006). The amplitude of this echo can be measured to estimate the size of the crack (Zangenehmadar, 2016). Furthermore, the use of flexural waves has also proven to provide insight into the prediction of longitudinal cracks, although the propagation of flexural waves is harder to initiate(Lowe & Cawley, 2006).

This method is generally used for above-ground pipes, in which a single probe can assess 30 m at a time (Royer, 2012). However, this condition assessment method has also been applied in buried pipes, when the desired inspection area is small, but, due to the rapid dissipation of the ultrasonic wave in the surrounding environment, this method has a limited application (Liu & Kleiner, 2012).

2.4.12 Discrete ultrasound

Discrete ultrasound condition assessment methods transmit a high-frequency wave that propagates through the inspected pipe and records echoes. Discrete ultrasound technology only inspects the area near the transducer (Royer, 2012). The transducer is installed in contact with the pipe wall and is moved along to perform the assessment. The advantage of this method is its sensitivity to the surface and subsurface corrosion that is harder to detect in the previous ultrasonic methods. However, in this method, the pipe needs to be exposed. Furthermore, the inspected pipe materials for ultrasonic methods need to be homogeneous in their composition (i.e., PVC and metallic pipes). If not, the interpretation of the phase signals is complex and it is very difficult to distinguish between a defect and a change of material (Zangenehmadar & Moselhi, 2015).

2.4.13 Summary of direct condition assessment methods

The previous sections described the main methods for direct condition assessment. A summary of the advantages and disadvantages of each method is presented in Table 2-4.

The methods that require excavation or access to the exterior of the pipe are less attractive, as a result of the additional cost added to the assessment, as well as risk of damaging the pipe during excavation works. Methods that are limited to only one type of material are more appropriate for water transmission system, rather than distribution networks. The reason behind this is the fact that transmission systems are homogeneous in terms of pipe material.

Regarding the difficulty of choosing a direct condition assessment method, Agarwal (2010) developed a framework to assist in the selection of the method used, referred as condition assessment selection tool (CAST). This framework classifies each assessment method with five distinct families of parameters that provide aid to choose the most adequate method of assessment. The framework facilitates the process of choosing an appropriate method by providing three indexes – total performance index, tangible cost index and intangible cost index – which allow to compare similar methods. However, the framework requires a survey of market expertise on each method before its application, which can require extensive research.

Overall, although a diverse portfolio of techniques may be available for direct condition assessment, the methods with the largest market share usually provide the most competitive price and, consequently, are the preferred choice. Currently, the most used assessment method is CCTV inspection and, therefore, is the most conventional choice for direct condition assessment (Covas *et al.*, 2018; Feeney *et al.*, 2009).

Table 2-4 - Summary of direct condition assessment methods

Assessment methods	Advantages	Disadvantages
Closed-circuit television (CCTV)	– Optimal price. – No need for service interruption. – Appropriate for all pipe materials.	– Results are subjective and dependent on the operator analysis. – Unfeasible of application in an entire network.
Laser scan technology	– Improvement of CCTV results. – 2D or 3D map of pipe interior.	– Only used in large diameter pipes. – High level of expertise to combine CCTV results with pipe profile. – Need for service interruption.
Magnetic flux leakage	– High accuracy data results. – Interior and exterior analysis.	– Limited to metallic pipes. – Limited to only large diameter and clean unlined pipes.
Remote field eddy currents	– No interruption of service required (possible). – Direct contact with pipe wall not required. – Results not effected by tubercules.	– Needs skilled technician to interpret results. – Some tools require pipe cleaning if tuberculation is excessive.
Broadband electromagnetic	– Does not require contact with pipe wall. – Can scan through pipe linings and insulation.	– Only measures thickness under sensor. – Pipe needs to be cleaned and emptied. – Slow and expensive method.
Ground-penetrating radar (GPR)	– No interruption of service required. – Underground mapping. – Trenchless method.	– Not frequently used in pipe condition assessment. – Dependent on soil conditions.
Ultra-wide pulsed radar system	– No interruption of service required. – Higher quality and resolution than GPR. – Applicable in all material types.	– Not frequently used in pipe condition assessment. – Restrictions on wave lengths used.
Sonar	– No interruption of service required. – Precise readings of pipe cross-section.	– Difficulty in recognizing longitudinal cracks. – High level of expertise required.
Smartball	– No interruption of service required. – Appropriate for large scale inspection (16 - 18-hour runtime). – High accuracy results.	– Inspection route uncontrollable.
Guided wave ultrasound	– Single probe can inspect 30 meters of pipe. – Inspection of hidden structures and under coatings, insultation.	– Access to the exterior of the pipe required.
Discrete ultrasound	– No interruption of service required. – Surface and subsurface inspection. – Instantaneous results.	– High skill level required. – Access to the exterior of the pipe.

2.5 Indirect condition assessment

Indirect condition assessment methods generally do not require as much field research or surveying to produce similar results to direct methods. Indirect methods are based on previously developed prediction models and known correlations between pipe properties and external factors to infer about the infrastructural condition of the pipe. The prediction models are often the quantitative relationships between pipe properties (i.e., physical, operation or environmental) and the corresponding pipe condition assessment index (Liu & Kleiner, 2014).

The main advantage of indirect condition assessment methods compared to direct methods is that inspection is not necessary in most cases. As a result, the interruption of the service is avoided as well as the cost associated (Ugarelli & Bruaset, 2013). Despite the many advantages of these methods, these are data-intensive and require preliminary data analysis to develop robust and reliable prediction models.

2.5.1 Overview of deterioration models

Indirect condition assessment methods predict the physical condition of the assets based on indirect variables. These variables also allow to estimate future condition of assets, as they estimate the deterioration rate overtime (Steven & Jenkins, 1999; Ugarelli *et al.*, 2010). Deterioration rate is defined as *"the amount of degradation of condition from its original condition"* experiences in a period of time (Maji & Jha, 2007) and can be interpreted as the opposite of asset performance, assuming that the maximum level of deterioration corresponds to the minimum level of performance (Gaspar, 2009; Matos, 2015). The decrease in asset performance can be evaluated in various forms, some of which are the deterioration curves.

Deterioration curves proposed by Shoet, Rosenfeld, *et al.*, (1999) show four types of behaviours depending on the deterioration mechanism at stake (Figure 2-8). A concave deterioration curve indicates the presence of biological factors, in which the initial phase of service the asset experiences has high levels of deterioration which gradually decrease over time (Figure 2-8a). A convex deterioration curve indicates the presence of physical and chemical factors which initially present little signs of deterioration, but the cumulative effects of these factors become gradually noticeable over time (Figure 2-8b). A linear deterioration curve indicates that deterioration agents, or influencing factors of physical condition, are constant throughout the asset service life (Figure 2-8c). Finally, the "S" deterioration curves indicate factors that change over time or the combination of the factors associated with convex and concave deterioration curves (Figure 2-8d) (Gaspar, 2009; Matos, 2015; Shoet *et al.*, 1999).

A study into the factors that affect the physical condition of pipes is presented in the following section to betted understand the behaviour of water distribution pipes during their service lives as well as the deterioration mechanisms that cause the shape of the deterioration curves.

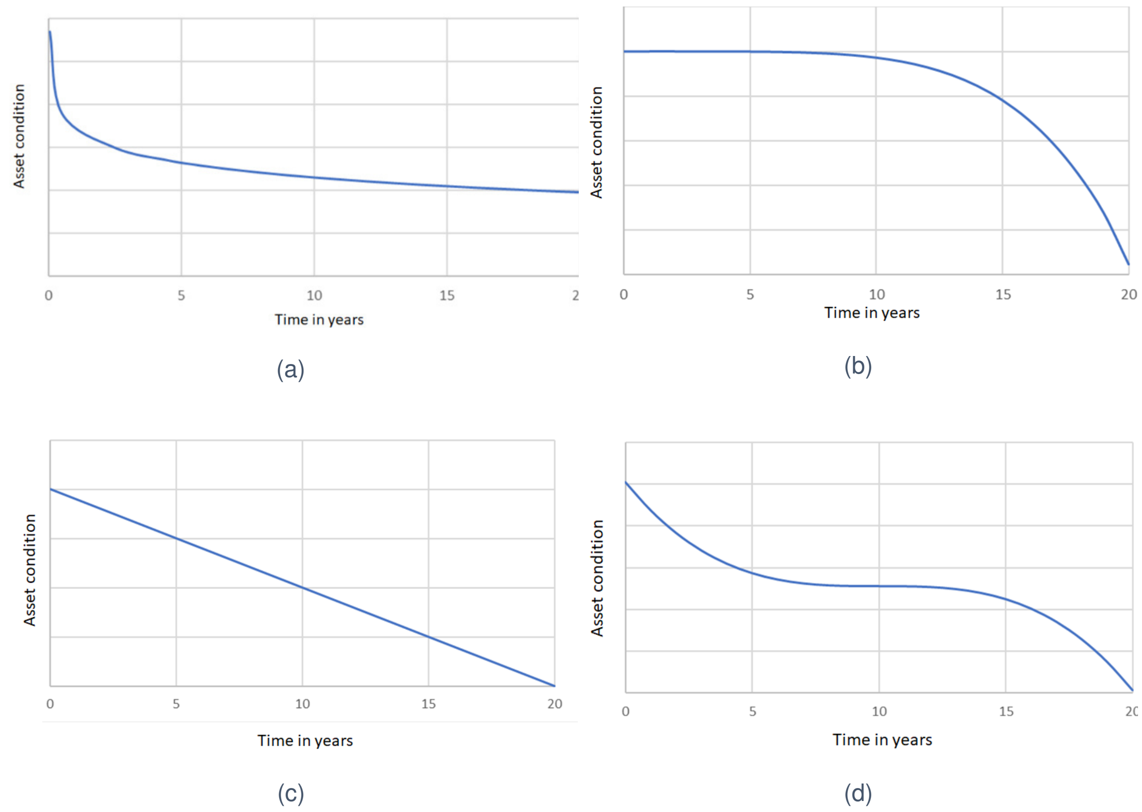


Figure 2-8 – Deterioration curves: (a) Concave deterioration curve; (b) Convex deterioration curve; (c) Linear deterioration curve; (d) “S” deterioration curve.

2.5.2 Factors affecting the physical condition of water distribution pipes

2.5.2.1 Overview

The factors that affect physical condition of buried pipes can be divided into (Al-Barqawi & Zayed, 2006a):

- **Physical factors** – factors that are related to intrinsic pipe properties, such as pipe material, age, diameter, wall thickness, pipe lining and coating, type of joints and manufacturing process.
- **Environmental factors** – factors that are related to the environment in which the pipe is located, such as soil pH, groundwater level, pipe location in relation to traffic, seismic activity, stray electric currents, trench backfill material, pipe bedding, installation practices and underground disturbances.
- **Operational factors** – factors that are related to the operational modes in which the pipe operates, such as water quality, water pressure, backflow potential, flow velocity, and operational and maintenance practices.

A single burst, as a result of increased pipe deterioration, can be attributed to a single factor or a combination of factors. While the factors that contribute to the deterioration of water distribution pipes are known, the relative influence each factor has is widely discussed (Rajani & Kleiner, 2001b). As a result, even though research has proven the existence of failure mechanisms, the uncertainty associated with the exact relationship factor is still the focus of many studies.

In a set of 15 peer-reviewed papers, focusing on the indirect assessment of the condition of water pipes, 42 different parameters are identified as explanatory factors of pipe condition, being the average number of influencing factors considered in each study approximately 7. The most frequently used factor is pipe nominal diameter, where 13 out of 15 studies used this parameter, followed by pipe age and type of material, which are considered in 11 and 9 studies, respectively. Overall, more than 75% of parameters is used by two or less studies, with 47% of the studied parameters being unique to a single study. The following section explores the main influencing factors, which have proven to have a clear effect on pipe condition.

2.5.2.2 Effect of nominal diameter on condition

A strong relationship between pipe failure and nominal diameter is supported by the literature (Barton *et al.*, 2019; Hu & Hubble, 2007; Kimutai *et al.*, 2015). The highest failure rate is reported to be in small diameter pipes (i.e., less than 200 mm), but this is also the most common range of diameters in distribution networks, as a result, conclusions need understood with parsimony (Bruaset & Sægrov, 2018).

Additionally, the failure mode of distribution pipes differs from small to large diameter pipes. While small diameter pipes are more vulnerable to beam failure, cumulative tuberculation, pressure instability and construction works in the surrounding environment (Al-Barqawi & Zayed, 2006b; Rajeev *et al.*, 2014a), the main failure mode of large diameter pipes is longitudinal cracking and shearing at the bell end of a pipe as a result of high water pressure and moment of inertia (Rajeev *et al.*, 2014b).

Figure 2-9 presents the failure modes of grey cast iron pipes according to diameter ranges (Makar, Desnoyers, *et al.*,2001). Figure 2-9(a) shows that, in the failure modes associated with small diameter pipes (i.e., less than 380 mm), the pipe on the left exhibits bell splitting at the top, cracking circumferentially in the middle, and the right pipe presents corrosion through a hole, as well as a chain created corrosion in the middle of the pipe with a blowout hole on the bottom. Figure 2-9(b) demonstrates the failure mode of spiral cracking, which can be associated with medium range diameter pipes (i.e., between 380-500 mm). Figure 2-9(c) presents the failure modes associated with large diameter pipes, where the left pipe presents longitudinal cracking and the right pipe presents bell shearing (Makar *et al.*, 2001).

Figure 2-9 – Failure modes by diameter in cast iron pipes: (a) Failure mode for small diameter pipes (< 380 mm); (b) Spiral mode of failure present in mid-diameter pipes (380 – 500 mm); (c) Failure mode for large diameter pipes (> 500 mm) (Makar *et al.*, 2001).

2.5.2.3 Effect of material on pipe condition

The service life of an asset is primarily defined by its material (Alegre & Covas, 2010). The reason behind this is the fact that each material presents a different durability which is directly associated with the estimated service life. The service life of water pipes can be increased or decreased depending on the operational and environmental factors that affect the deterioration mechanism (Ruchti, 2017; Ugarelli *et al.*, 2010) as well as O&M practices (Covas *et al.*, 2018).

In metallic pipes, the main cause for the decrease of pipe condition is the formation of corrosion pits in the outer pipe walls. In this specific form, the main contributing factor is the electrochemical environment (Rajani & Kleiner, 2001a). In grey cast iron pipes, graphitization typically occurs, which is the formation of iron oxide flakes on the pipe surface (SUEZ's, 2020).

On the other hand, Table 2-5 presents the reference values of established service lives for different pipe materials. These values were established through the judgement and consultation of experts in the urban water field for a reference in-use condition of assets. The service life values are only for newly installed water pipes and not for relined pipes or any other form of rehabilitation carried out. Nonetheless, NSW(2014) proposes an average of 50 years for the service life of relined pipes and 80 years for new pipes.

The service lives presented in Table 2-5 have different values depending on the source. Firstly, it should be noted that the asset accounting life is defined as the period of tax amortization (Covas *et al.*, 2018). This type of asset life serves an economical purpose that is different to the economical purpose of the established technical service lives. On the other hand, the existing changes in the technical guidelines regarding the service life by materials is the result of the evolution in manufacturing processes, which have improved the quality and durability of pipes as well as the improvement of O&M practices by water utilities (Covas *et al.*, 2018; Rezaei *et al.*, 2015). However, the established service lives of water distribution pipes tends to be higher in Central Europe, Australia and the USA due to higher O&M practice levels in these countries (Covas *et al.*, 2018).

Table 2-5 – Established pipe service life by material.

Pipe material	**Accounting service life (according to article 3º DR 25/2009)**	**Technical service lives - Portugal (*Alegre and Covas, 2010)***	**Technical service lives - Portugal** (Covas *et al*., 2018)	**Technical service lives - USA** (American Water Works Service Co., 2002)
Asbestos cement (AC)	16	30	40	-
High density polyethylene (HDPE)	-	45	50	50
Cast Iron (CI) and Steel	25	40	60	100
Concrete	20	60	50	-
Polyvinyl chloride (PVC)	-	50	40	50-100

2.5.2.4 Effect of pipe age on pipe condition

In general, age is a common parameter used for deciding on pipe maintenance and rehabilitation, although many authors reject this approach, because it does not account for the operation and environmental factors variability and other contributing factors (Andreou *et al*., 1987; Babovic *et al*., 2002). Another example of this application, is the case of water utilities in France estimating their pipe renewal requirements based on the notion that all pipes must be replaced when they reach a certain age, predefined for each material (Renaud *et al*., 2014). Still, pipe age has shown to be the main contributing factor to water pipe condition (Al-Barqawi and Zayed, 2006a; Berardi, Giustolisi, *et al*., 2008). However, this factor is not appropriate to predict pipe burst rates (Babovic *et al*., 2002).

Additionally, authors have used pipe age as a proxy parameter when calculating pipe burst rates in order to address the continuous physical and environmental factors that affect the pipe through service life (Hu & Hubble, 2007). Hu and Hubble (2007) reviewed the factors affecting the failure in asbestos-cement pipes and stated that pipe age as a proxy factor to analyse the deterioration caused by static factors (i.e., pipe material, pipe diameter, type of soil). As a result, the authors could isolate the static effect of deterioration and investigate the effects of rainfall deficit and freezing climates on deterioration rates. The study concluded that rainfall deficit caused a significant shrinkage of the clay soil which caused additional stress on the distribution pipes. Furthermore, the study also found that that the occurrence of an initial failure in a certain location increased the potential for future failures, likely as a result of disturbances in the soil associated with pipe repair.

The effect that age has on pipe condition is directly correlated with deterioration effects. This is due to the longer pipes are in use, the more exposure they are to internal and external agents of deterioration (i.e., influencing factors) (National Research Council, 2003). Consequently, age is an important factor to consider when developing predictive models of pipe deterioration (Berardi *et al.*, 2008; Ennaouri & Fuamba, 2013).

2.5.2.5 Effect of operational pressure on pipe condition

In the Portuguese legislation, DL nº 194/2009, the legal responsibilities of water utilities are defined, namely the responsibilities in safeguarding the integrity of water supply systems, which are established in article 71 . The legislation states that water utilities are responsible to mitigate abnormal deterioration of water pipes caused by excessive pressure or by sudden pressure variations. The relevancy of this information for this research is that it proves that the correlation between accelerated deterioration rates and operational pressure is not only already established, but significant enough to become legally stated.

Also, the relationship between the reduction of operational pressure and burst rate is clear (Pearson *et al.*, 2005). The unsteady hydraulic conditions in water pipes can initiate the growth of micro-fractures and corrosion fractures throughout the network by slowly decreasing the pipe tensile strength enough to cause pipe failure. Not only this, but the growth of micro-fractures and existing fractures can be increased by excessive operational water pressure (Rezaei *et al.*, 2015).

Nevertheless, depending on the type of fracture exhibited, the operating pressure can have a larger or smaller influence on the evolution of the fracture. For example, while fracture growth rates are linearly affected by the increase in pressure, rounded holes have shown to be less affected by these effects. On the other hand, longitudinal cracks expand more with an increase in pressure and circumferential cracks showed to have a middle-range value of growth in comparison to the other two types of fractures (Cassa *et al.*, 2010).

2.5.2.6 Effect of soil properties on pipe condition

The electrochemical environment surrounding pipelines is one of the sources of environmental deterioration mechanisms. Consequently, the relevance of soil properties, when discussing the contributing factors for pipe deterioration becomes relevant.

The soil properties that influence pipe condition can either be direct contributors, such as the case of soil contamination or the presence of sulphates, or can be indirect contributors, increasing the effect of direct factors. A summary of soil properties and their contribution to the deterioration of pipes and their classification as direct and indirect contributing factors is presented in Table 2-6.

The soil resistivity is a relevant parameter in soil properties that is influenced by the temperature, the moisture content, the naive type of soil and the presence of salts (Csanyi, 2015). The evaluation of soil resistivity allows for estimation of corrosion rates in the pipes installation. In general, low soil resistivity indicates high soil corrosiveness, which result in high deterioration rates (Arbabi, 2005).

Table 2-6 – Summary of soil properties and their effects on pipe condition (Wasim, Shoaib, *et al.*, 2018; Asset Management Condition Assessment Techniques, 2002; Liu & Kleiner, 2012).

Soil property	Deterioration mechanism	Direct/Indirect
Moisture content	Acts as an electrolytic corrosion vehicle for the pipe-soil interaction	Direct
Sulphate salts	React to cementitious materials forming ettringite and gypsum	Direct
Soil aeration	The presence of oxygen in the soil promotes the microbial induced corrosion	Indirect
pH	Imbalance of hydrogen ions promotes the ferrous reaction (either pH<4 or pH>8)	Direct
Chloride content	Aid in electrolytic corrosion as acting as a cathode	Direct
Temperature of soil	Effects pH and moisture content in the soil	Indirect
Shrink/swell capacity	High shrink/swell capacities induce stress cycles	Direct
Contaminates	Increase pH and promote electrolytic corrosion	Direct
Soil compaction	Increases or decreases shrink/swell capacity	Indirect
Microbial activity	Acts as an electrolytic corrosion vehicle for the pipe-soil interaction	Direct

2.5.2.7 Effect of water quality on pipe condition

The aging effects of water distribution pipes can have a negative impact on water quality (American Water Works Service Co., 2002), as they can increase the presence of solid matter in the water system and, consequently, can potentially be a hazard for the serviced population (Ugarelli *et al.*, 2010). However, the effect water quality has on pipe condition is equally evident. The parameters inherent to the definition of the water quality in the distribution system can be physical (e.g., turbidity, temperature, colour, solids), chemical (e.g., pH, chlorine residual, hardness, dissolved oxygen) or biological (e.g., bacteria, algae, viruses) (Hassan Omer, 2020). The combination of different parameters at different intensities can have a positive or negative effect on pipe deterioration rates. An example is the effect of pH levels on cement-lined and cement-based water distribution pipes. Alkalin water (i.e., pH<7) is associated with leaching, caused by the dissolution of calcium hydroxide decreasing the structural integrity of water pipes resulting in increased deterioration rates (World Health Organization, 2014).

The events associated with the humanitarian crisis in the city of Flint, Michigan, were also related to water quality parameters, namely chlorine levels and phosphate inhibitors. The addition of orthophosphate, a type of phosphate inhibitor, often used in water distribution systems, allows for the formation of a passive layer of protection from corrosion (Selvakumar *et al.*, 2013). Following the switch from the Detroit water system to the Flint River, orthophosphate was not added during the water treatment and, as a result, the passive layer of protection started to decay (Torrice, 2016). Additionally, the Flint river presented a lower pH level and a naturally higher chlorine concentration estimated to be a result of runoff from road de-icing salt (Green, 2019). Consequently, the higher concentration of chlorine reacted with the lead pipes causing lead and iron levels in the water to increase. Furthermore, the chlorine disinfectant added for bacterial control reacted with the lead and iron pipes further increasing the levels of these components in water and rendering the disinfecting compounds useless (Woszczynski *et al.*, 2013). As a result, high levels of fecal coliform bacteria (i.e., e-coil) were detected causing an initial public health warning. However, the bacterial outbreaks caused the water utility to increase chlorine disinfectant dosage, which further escalated the corrosion rates of iron and lead pipes causing toxic levels of metals to be present in the water system (Masten *et al.*, 2016). These events began on April 25th, 2014, and the water source was switched back to the Detroit water system on October 16th, 2015 with extra concentration of phosphate inhibitors (Emergency Administirve Order, 2016). However, along with the deaths caused by the bacterial outbreak and permanent health damages to the population, distrusts in local authorities remained, while extensive rehabilitation efforts of the city's water distribution systems are underway (Carmondy, 2017).

2.5.3 Infrastructure value index

The variety of influencing factors described in the previous section illustrate the physical mechanisms and contributing factors for the deterioration process. The evaluation of the asset condition requires the measurement of key performance indicators (KPI) that translate asset performance. The most notable KPI is infrastructure value index, since it requires minimal data and delivers reliable results (Alegre *et al.*, 2014). The following section will explore indirect condition assessment methods with incremental complexity in their conception.

The infrastructure value index (IVI) developed by Alegre (2008) is the ratio between the current value of an infrastructure and the replacement cost to the modern equivalent and applies to infrastructures composed of different assets (Alegre *et al.*, 2014). Which can provide a general perspective of the condition of the infrastructure (Cabrera Rochera *et al.*, 2019), as well as provide a useful tool to set targets for infrastructural sustainability (Cabral *et al.*, 2019). IVI is described by:

$$IVI(t) = \frac{\sum_{i=1}^{N} \left(cs_{i,t} \cdot \frac{vr_{i,t}}{vu_i} \right)}{\sum_{i=1}^{N} cs_{i,t}} \tag{2.1}$$

in which t is time (years); $IVI(t)$ is the infrastructure value index at a time t (-); $cs_{i,t}$ is the cost of replacement (€) of the asset i at the time t; $vr_{i,t}$ is the remaining service life (years) of the asset i at the time t; vu_i is the service life (years) of the asset i. When applied to a single asset, this index represents the percentage of remaining service life (or residual life).

The IVI for a set of assets, or the percentage of residual life for one asset, has proven to be a useful tool for the long-term planning of linear (i.e., water distribution pipes, water transmission pipes) and vertical assets (i.e., water storage tanks, pumping stations, treatment plants) (Alegre *et al.*, 2014) and adequate for determining infrastructural integrity (Alegre & Covas, 2010). The use of this index is relevant for (Alegre, 2008):

- Defining in planning and investment strategies in asset management plans.

- Having a general perception of the network at different times.
- Establishing goals for concession contracts, regulators or internal assessment by water utilities.
- Establishing guideline policies for asset rehabilitation at the regional or national scope.

The calculation of IVI can be carried out using one of two approaches:

a) *Asset-oriented*: calculations are based on the service life of assets, linear depreciation curves and the replacement costs for each category of assets.
b) *Service-oriented*: calculations are based on the performance of functional unit of infrastructure.

The service-oriented approach acknowledges that the value of infrastructures is directly related to the quality of service provided and that useful life should be an output of performance and not an input for planning. However, the results provided by an asset-oriented have shown to be less subjective because of the avoidance of defining quality of service benchmarks (Alegre *et al.*, 2014).

Additionally, when interpreting the IVI results, if all assets in the infrastructure have the same replacement cost, the IVI values can be interpreted as the residual life. Or in a more realistic scenario, the IVI values would be interpreted as a percentage of the weighted average of residual lives (Alegre *et al.*, 2014).

IVI calculation can be performed for both vertical assets (i.e., pump-stations, water tanks) and linear assets (i.e., transmission or distribution pipes). In the case of linear assets, to correctly calculate the IVI value, Alegre (2008) has proposed a methodology that firstly requires the collection of asset-related data, which include information regarding properties of assets (e.g., configuration, material, nominal diameter) as a means to understand the system and the functionality of the assets. This step might also require information related hydraulic properties (e.g., Hazen-Williams coefficient, pressure head, pressure variation), as well as information on the cost of installation and any rehabilitation works that might have been completed. The information collected from this first step is needed to estimate the cost of replacement of assets, which requires the calculation of the modern equivalent asset that provides the same level, or similar level of service potential (Adorno *et al.*, 2018).

The IVI results for a network in which the water utility has good maintenance practices should vary between 0.4 and 0.6. Higher values might indicate younger networks, old networks experiencing expansion, or old networks experiencing over-investment in rehabilitation efforts. Low IVI values can be correlated to poor maintenance practices and low rehabilitation efforts (Alegre *et al.*, 2014).

The IVI scale (Table 2-7) is similar to the condition grading scales previously described but has the advantage of not requiring any additional information collected in the asset inspection.

Table 2-6 – Infrastructure value index description proposed by Alegre (2008).

Infrastructure value index	Description
1 – 0.6	Young network or in an expansion phase
0.6 – 0.4	Stabilized network
0 – 0.4	Aged network

When analysing the IVI of a distribution network, the expected value is generally above 0.4, this is the result of preventive maintenance and rehabilitation efforts being generally carried out and (Jin Jun *et al.*, 2020). However, IVI values can be unbalanced, meaning that, if an infrastructure has assets that are both new and old, the weight of each can present a bias that skews IVI results, allowing for wrong final assessments.

To prevent this problem, Cabrera Rochera, *et al.* (2019) have proposed the infrastructure degradation index (IDI) to complement the IVI analysis. IDI assesses the remaining useful life of the network weighted by the length of each pipe, described by:

$$IDI(t) = \frac{\sum_{i=1}^{N} L_{i,t} \times vr_{i,t}}{\sum_{i=1}^{N} L_{i,t}} \tag{2.2}$$

where t is time (years); $IDI(t)$ is the infrastructure deterioration index in the year t; $L_{i,t}$ is the length of the pipe i in the year t; $vr_{i,t}$ is the remaining service life (years) of the asset i at the time t.

Similar to IVI, IDI varies from 0 to 1. However, IDI differs from IVI as the interpretation of the results obtained allow to estimate what percentage of the pipe network is reaching the end of its service life, since it applies only to pipes. Furthermore, IDI is expressed in years, allowing water utilities to assess how much time is needed until all the pipes of the network reach the end of their service lives.

2.5.4 Deterministic models

To estimate the physical condition of assets the relationship between the physical, environmental, and operational properties can rely on deterministic modelling approaches. A deterministic approach can be defined as a modelling attempt where there is no randomness in the input variables (Lawson & Marion, 2008). Consequently, the relationship between the input properties and the resulting deterioration mechanisms are unique. The main advantage of deterministic models is that they can be easily understood and applied in asset condition assessment (Matos, 2015). However, the use of deterministic approaches is considered inappropriate in scenarios where the risk of an event is a relevant criterion for decision making (e.g., seismic risk, vulnerability studies). In this case, deterministic approaches are unappropriated since they are highly reliant on initial conditions and do not consider the probability of variants (Cassiolato *et al.*, 2021; Uusitalo *et al.*, 2015).

The use of a deterministic approach is valid when the relationship between the physical condition of assets and a set group of variable is perceived to be certain (St. Clair & Sinha, 2012). Two methodologies have been used in the deterministic modelling for condition assessment: empirical and mechanistic. In both methodologies, the prediction models are developed through regression analysis and can be combined with the opinions of experts.

In empirical modelling, models are based on the assumption that the relationship between the rate of failure and characteristics is non-random for water distribution pipes (Zangenehmadar, 2016). The chosen characteristics to develop these relationships can be physical, environmental or operational and can be used in any combination that aim to better predict pipe condition or associated parameter (Rajani & Kleiner, 2001b). The application of an empirical approach is more appropriated when applied to homogeneous groups of pipes, also known as cohorts, by respecting their primary influencing factors (Ugarelli & Bruaset, 2013). Moreover, the aggregation of pipes into cohorts has to follow a set of guidelines to obtain representative results (Wood & Lence, 2009). The guidelines vary in the literature, but the common principle behind is that the cohorts must respect their overall characteristics (Rajani & Kleiner, 2001b). An example of this division into groups is proposed by Ugrarelli and Bruaset (2013), which used the following rules to aggregate pipes into cohorts:

i. The cohorts must be small enough to be considered uniform.
ii. The cohorts must be large enough to provide statistically relevant information.

Finally, an empirical deterministic modelling approach must be time-dependent, meaning that considerations about future deterioration should also be made (Ruchti, 2017).

Similarly, deterministic mechanistic models aim to estimate the remaining service life of pipes through the known failure mechanisms (St. Clair & Sinha, 2012), as well as to describe the structure of events that result in the deterioration of the assets (Lawson & Marion, 2008). The advantage of estimating the remaining service life is that the output information of these models can be directly coordinated and used at optimal times of preventive maintenance interventions (Zangenehmadar, 2016). While these models have shown to be more robust (Rajani & Kleiner, 2001b), their application has been limited to one single material type at once. Since different materials have different deterioration mechanisms, deterministic mechanistic condition assessment models have difficulties in combining deterioration rates that result in a final condition assessment classification. Additionally, in some mechanistic models, the cost of acquiring the necessary information to evaluate the influence of some of the parameters can be so high, leading the models that can only be applied in highly relevant or critical water pipes (Wood & Lence, 2009).

An example of the application of combined deterministic and mechanistic models is the work done by Rajani and Makar(2000). The study aimed to estimate the remaining service life for grey cast iron water pipes. In this study, the author used a total of 29 parameters to estimate service life, which included the measurement of corrosion pit size obtained through maintenance works or the indirect assessment of surrounding soil properties to predict corrosion rates. The measurement of corrosion was carried out twice in order to evaluate the stress-state and to obtain a growth rate. Following the assessment of the stress-state, results were compared to a safety factor defined by the water utility, which acted as a threshold for the repair or replacement of the pipe. If the pipe did not require replacement, the resulting information was used along with the relationship between residual tensile to determine residual service life. Although this study was extensive and has proven to have good results, the application of this type of deterministic approach by any water utility is unrealistic. The data collection costs required for this type of assessment are too high for it to be considered an option (Thomson & Royer, 2013).

2.5.5 Statistical and probabilistic models

Statistical and probabilistic models, also known as stochastic models, are those which consider variable parameters typically described by statistical distributions (Martins, 2011). Unlike deterministic models, these models account for uncertainties in input data and use probability theory to approach these uncertainties (Oliver Stegle, 2020). For condition assessment, this is an important aspect since deterioration is inherently random in occurrence and development.

However, stochastic models can also be mechanistic or empirical, as shown in Figure 2-10. The difference between the two modelling approaches is the level of hierarchical understanding models are based on.

Hierarchical level of consideration

Randomness consideration

Modelling types	Empirical	Mechanistic
Deterministic	Prediction of cattle growth from **regression relationship** with feed intake	Planetary motion, based on Newtonian mechanics **(Differential equations)**
Stochastic	Analysis of **variance** of variety yields over sites and years	Genetics of small population based on mendelian inheritance **(Probabilistic equations)**

Figure 2-10 – Differences and examples of model classification (Lawson and Marion, 2008).

Similar to deterministic mechanistic models, statistical approaches also aim to predict the remaining service life of assets (St. Clair & Sinha, 2012). However, statistical models do not discard the uncertainty associated with output variables. A common form to consider this uncertainty is to use Bayesian belief networks (Thomson & Royer, 2013). These networks can handle the probabilistic relationship between variables as well as of missing data (Kabir *et al.*, 2015).On the other hand, probabilistic models aim to predict the probability of a certain event, namely the probability of pipe bursts, similar to deterministic empirical models (St. Clair & Sinha, 2012).

The most notorious probabilistic model is the *bathtub curve* use to describe the pipe failure rate over time. The bathtub curve, shown in Figure 2-11, describes the rate of occurrence of failure (ROCOF), or hazard rate, with the lifetime of a buried water pipe (Rajani & Kleiner, 2001b). The curve demonstrates three distinct phases:

1. **Burn-in phase,** characterized by a high failure rate in the initial years of service, which can be attributed to defective installation.
2. **In-usage phase**, which is considered the service life period of operation of the pipe and is characterized by a constant failure rate associated with random occurrences, such as earth movements, extraordinary loading or third-party interference (Singh & Adachi, 2013).
3. **Wear-out phase,** where the signs of deterioration start to show in the form of pipe bursts and the effect is an increase in the failure rate (Rajani & Kleiner, 2001b).

An example of the use of the bathtub curve is in prioritizing interventions, as developed by Singh and Adachi (2013). In this study, a pipe prioritization process in terms of probability of burst was carried out. The study modelled the failure rate for each material type and used it to show that ductile iron pipes had the highest failure rates for pipes of 40-60 years old, while PVC and concrete cylinder pipes had the lowest failure rate. The guidelines followed by the study to aggregate pipes in cohorts were materials and age. The study was successful in modelling the bathtub curve for pipes of cast iron, ductile iron and asbestos concrete, although not so successful in PVC and concrete cylinder pipes.

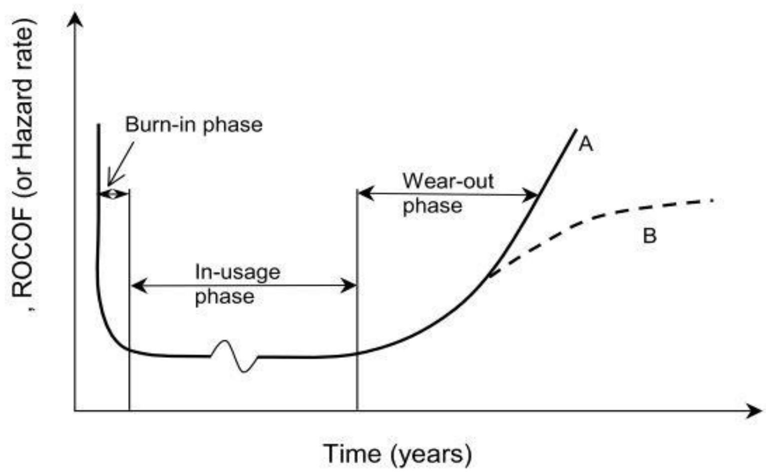


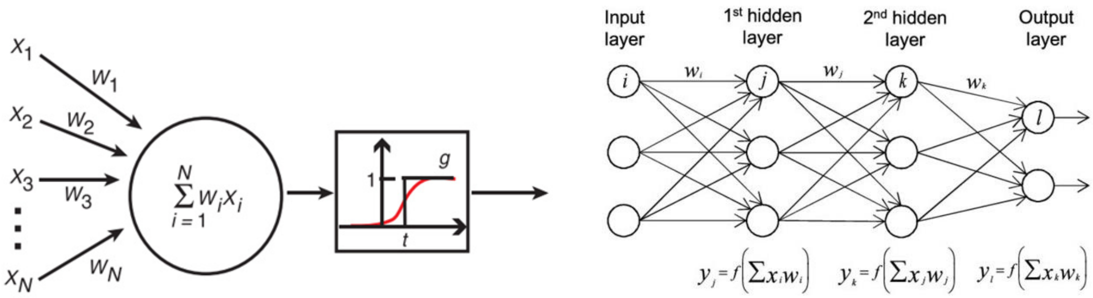
Figure 2-11 – Bathtub curve of the life cycle of a buried pipe (Rajani & Kleiner, 2001b).

2.5.6 Artificial neural networks

Artificial neural networks (ANN's) are biologically inspired computational models (Majumder, 2015). The concept is the simulation of a network of algorithms, known as neurons, which work collectively to obtain useful information (Figure 2-12). ANN is a type of model for machine learning and has recently become appealing as an alternative to traditional regression analysis (Abiodun *et al*., 2018). Each neuron works as a threshold unit, receiving values from other units and/or external sources and summing the values, to either achieve the threshold value and to output one, or not achieving the threshold value and output a 0. Then, the following set neurons receive the threshold output and compute their algorithm until all values culminate in one or more outputs (Krogh, 2008). An illustration of a working neuron as well as a representation of the inputs and hidden layers are presented in Figure 2-12.

ANN have become increasingly popular in recent years (Abiodun *et al*., 2018) due to several advantages that have proven to have, namely (Mahanta, 2017):

- The ability to process and to model non-linear and complex relationships between variables, allowing to model real-world scenarios where these relationships are required.
- The better modelling of heteroskedasticity because of the absence of input variable restrictions or distribution formats.
- The ability to infer upon previous relationships, allowing observe unseen patterns.



(a) (b)

Figure 2-12 – Representation of the functionality of an ANN: (a) function of an individual neuron (Krogh, 2008); and (b) representation of a connected neural network (Vieira *et al*., 2017).

In the field of condition assessment, artificial neural network allow modelling of the physical deterioration mechanisms as well as to overcome the difficulties associated with statistical and probabilistic modelling approaches when dealing with uncertain associations between influencing factors and deterioration mechanisms (Al-Barqawi & Zayed, 2006b).

Several studies have used ANN to predict the physical condition of networks, namely Al-Barqawi and Zayed, (2006b). In this study, the author collected information on physical (i.e., pipe age, pipe material and pipe diameter), environmental (i.e., type of soil, type of road surface and pipe cover) and operational properties (i.e., number of breaks, Hazen-Williams coefficient) to use as an input variables for condition assessment. After data collection the model was trained, meaning that, a supervised ANN, where input and output variables were known, used a backpropagation algorithm to develop the associated weights of each neuron and obtain a working ANN. The ANN was trained using a randomly generated portion of the acquired dataset and validated against the remaining data set. The results obtained were considered robust, with 71.7% of results being within 5% of actual values, and 91.7% within 10% of actual values. Furthermore, the weights associated with each neuron are physically significant, meaning that the influence that each factor had on the condition of the pipeline was able to be quantified for the analysed system.

Finally, it is important to highlight that, although this modelling approach is very popular and effective, two key aspects make it very difficult for water utilities to use. Firstly, ANN needs to be trained using known input and output variables, which requires a considerable set of data to train and test the ANN and, secondly, although, the development of ANN models requires specific mathematic and programming knowledge which water utility engineers do not have.

2.5.7 Heuristic models

In the field of psychology, heuristics is an approach of problem-solving where personal experience is a key contributor to the decision making, as well as a fundamental shortcut for streamlining decision-making processes when there are many decision or explanatory variables (Dale, 2015). When applying a heuristic approach to modelling strategies biases may occur in the form of overestimation or underestimation (Blum & Roli, 2003). However, the effectiveness of these types of approaches (i.e., trial and error, educated guessing, rule of thumb), as well as meta-heuristic approaches, such as evolutionary computation, simulated annealing, swarm particle optimization, have all proven to be successful when other approaches fail (Fukuyama, 2008).

In condition assessment methods the use of heuristic modelling is frequently associated with situations where the infrastructural problems are not well understood (St. Clair & Sinha, 2012). The use of heuristic models is needed in situations where data are limited and the application of more complex methods are unfeasible (Zangenehmadar, 2016).

The combination of an ANN with an analytic hierarchy process (AHP) (i.e., a form of heuristic analysis) to estimate beforehand the condition of the assets can be an alternative to improve confidence in results and overcome limitations in other modelling approaches. An example of this approach is the work developed by Al-Barqawi & Zayed (2008) that integrated an AHP with an ANN to reduce data requirement costs and obtained very good results.

2.5.8 Fuzzy logic models

Fuzzy logic can be characterized as multi-valued logic, meaning that they allow for larger number of truth values (Gottwald, 2020). The specificity of fuzzy logic datasets allows the adoption of any value between 0 and 1. The consequence of this multi-valued logic is that it grants statements to be categorized accordingly to the degree to certainty associated with each variable (Novák *et al.*, 2013). Fuzzy logic is directly opposed to Boolean logic, where statements can either be true (1) or false (0) (Zohuri & Moghaddam, 2017). Additionally, it should be noted that the steps in common in models that use fuzzy logic are described in Figure 2-13. The first step is the input of crisp data, which can be defined as the formal classification in which data is originally stored (Liang *et al.*, 2006). Then into a fuzzifier, which transforms the data to one of the logical values between 0 and 1 using membership functions, working as a measurement of influence each data entry has with its attributes. Then inference takes place, which uses a series of pre-established rules, obtained either through ANN, expert experience, or intuition to make decisions regarding the data. Following the inference, the data is then defuzzied into its original structure (Kayakan & Khanesar, 2016).

The application of this type of logic can be used in condition assessment for predicting the risk of pipe breaks and estimating the remaining useful life (Fares & Zayed, 2010; Tavakoli, 2018). Tavakoli (2018) combined a fuzzy logic system with ANN, resulting in an adaptive neuro-fuzzy inference system. The use of ANN allowed for the easy manipulation of the weight in the membership functions, improving overall results.

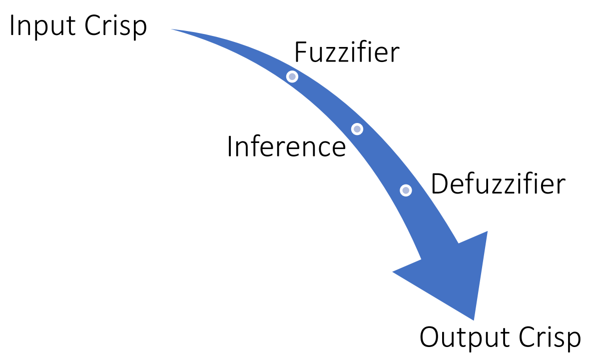


Figure 2-13 – Operational steps in a fuzzy-logic system adapted from (Kayakan & Khanesar, 2016).

2.5.9 Summary of indirect condition assessment methods

A summary of the advantages and disadvantages of the application of each indirect condition assessment method is presented in Table 2-7. Indirect condition assessment methods range in complexity and in data requirements. Elementary methods can yield accurate results and provide a useful basis for decision making. However, these methods do not often consider the arbitrary nature of deterioration mechanisms and, consequently, become unreliable. On the other hand, more complex methods of indirect condition assessments, such as fuzzy-logic, ANN or some stochastic methods, can overcome the dependencies of initial conditions, but can have rigid data requirements and require advanced knowledge on data processing and analysis making them difficult to use by water utilities.

Table 2-7 – Summary of advantages and disadvantages of indirect condition assessment methods.

Assessment method	Advantages	Disadvantages
Infrastructure value index	• Limited data requirements. • Easy comprehension. • Easy application.	• Highly dependent on initial considerations. • Revaluation with new information does not yield more accurate results.
Deterministic models	• Limited data requirements. • Easy comprehension. • Easy application.	• Do not consider the inherent randomness of deterioration mechanisms. • Too simple and misrepresentative. • Highly dependent on initial considerations.
Statistical and Probabilistic models (Stochastic approaches)	• Consider randomness of deterioration mechanisms. • Can portray complex relationships between variables.	• Interpretation can be often difficult for decision making.
Artificial neural networks	• High level of accuracy. • New data improves future results.	• Require computational know-how. • Difficult application. • Require extensive data.
Heuristic models	• Easy comprehension. • Easy application. • Allows to bridge gaps in other modelling approaches.	• Subjective nature. • Revaluation with new information does not yield more accurate results.
Fuzzy logic models	• Accurate results. • Randomness in the data is considered. • Combination with ANN for more robust results. • Can portray complex relationships between variables.	• Require computational know-how. • Difficult application.

2.6 Summary and conclusions

The current chapter presents a literature review on the most relevant aspects of condition assessment. The objective of the chapter is to give the reader a broad understanding of the state-of-the-art of condition assessment in water distribution systems.

While direct condition assessment methods are seen to be more reliable and have proven to profusely the structural integrity of assets, the application of such methods is often unfeasible due it their costly and labour-intensive nature. Furthermore, direct condition assessment methods are often too reliant on the experience of operators and can produce biased results. Consequently, the application of a direct condition assessment method as a form to evaluate the physical condition of assets, in a broader asset management plan, may not be appropriate and, therefore, is not the scope of current research.

On the other hand, the development and application of indirect condition assessment methods seem to be a good alternative for an asset management plan when the direct methods are not viable. However, these methods have also data requirements and application costs that need to be considered. The first step in the application of these methods is to evaluate the available information on the pipe network as well as the human resources in water utility to process and analyse those data.

Similar to the case of direct condition assessment methods, in indirect methods, data collection is often the critical part, thus, the use of mechanistic approaches to estimate the effects of deterioration mechanisms on pipes can be tempting but unrealistic. Hence, empirical approaches are recommendable in the case of limited available data.

Finally, the chosen modelling method to estimate either the pipe remaining service life, or the pipe failure rate, is highly dependent on the "know-how" and the available data of the utility. Therefore, in the case of water utilities with less infrastructural and operational knowledge, which is the case of most utilities in Portugal, the option of computationally intensive modelling approaches, such as fuzzy logic models and ANN, should be avoided, and deterministic and empirical methods should be considered an option.

The current indirect methods and methodologies for condition assessment of water distribution systems have become increasingly more accurate and robust. However, the application of these methods did not had the same progress. Therefore, the current thesis aims to address this gap of knowledge and to propose a methodology that can be used to estimate the physical condition of small diameter water distribution pipes.

3 Proposed methodology

3.1 Introduction

The current chapter presents the proposed methodology to assess the condition of water distribution pipes combining a deterministic and a heuristic approach. The deterministic approach describes the physical condition of buried pipes based on the relationship between the intrinsic network properties and the recorded failure history using linear and polynomial regression models. The heuristic approach uses a weighted sum model, whose weights are established by a survey carried out on urban water experts to quantify the importance of network properties and external factors on pipe deterioration. Both approaches should be complementary used when applying the proposed methodology in order to improve the robustness of results through the comparison of obtained results.

This chapter presents the general methodology, the detailed description of each approach and the respective assumptions considered, and the statistical analysis required for the deterministic model development.

3.2 General methodology

The proposed methodology for the physical condition assessment of water distribution networks is a four-step procedure (Figure 3-1). The application of this methodology aims to assess the physical condition of water distribution networks pipes without the need for inspection or extensive field survey. The application of this methodology also aims to contribute to the improvement of resource allocation and investment planning, resulting in the enhancement of the physical sustainability and integrity of water distribution networks.

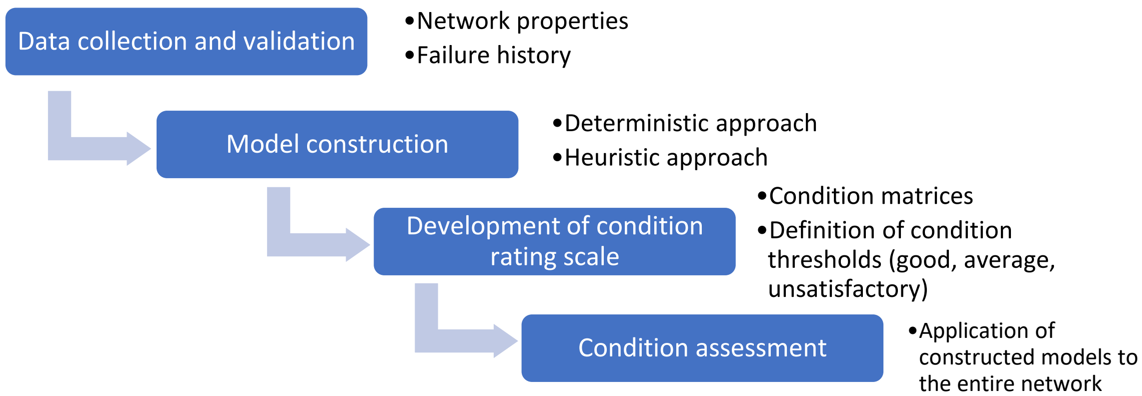


Figure 3-1 – Proposed methodology for the physical condition assessment of water distribution pipes.

The first step of the methodology consists of the collection of data related to the pipe network characteristics (e.g., network design, pipe location, pipe material) and failure history variables (e.g., number, location and occurrence date of bursts). In this step, the collected data needs to be validated for inconsistencies and discrepancies which requires the characterization of the assessed network.

The second step comprehends the construction of assessment models in order to evaluate the condition of buried pipes. This involves the application of a deterministic approach to develop linear or polynomial models that estimate the failure rate of a distribution pipe based on the pipes characteristics, as well as the application of a heuristic approach that allows to obtain a non-dimensional condition variable based on the results of expert surveys.

The third step consists of the development of a condition rating scale with three condition categories: good, average and unsatisfactory. A condition assessment matrix that combines the results of developed models taking into account the reference values of the performance indicator infrastructure value index is developed.

In the fourth step, the developed models are applied to a case study to assess the physical condition of several distribution pipes using the developed condition rating scale. The following sections will describe in detail each step of the proposed methodology and the respective considered assumptions.

3.3 Step 1 – Data collection and validation

The first step of the proposed methodology involves the collection of information regarding the factors that influence the physical condition of water pipes (e.g., pipe age, operating pressure, nominal diameter) and pipes failure history (e.g., date of burst, location), referred to as independent and dependent variables, respectively, in the model construction.

The collection of the largest and the most diverse number of factors is of utmost importance for a better characterization of the distribution network. In this study, a collection of 13 peer-reviewed papers were analysed to compile a list of the most relevant factors on the deterioration of water distribution pipes (Achim *et al*., 2007; Al-Barqawi & Zayed, 2006b, 2006a; Babovic *et al*., 2002; Berardi *et al*., 2008; Clark *et al*., 2010; Fahmy & Moselhi, 2009; Rajani & Makar, 2000; Tavakoli, 2018; Tuhovčák *et al*., 2014; Wang *et al*., 2009; Wood & Lence, 2009; Zangenehmadar & Moselhi, 2015), as described in section 2.5.2. The most important factors that affect the deterioration of water pipes can be grouped in four main categories: physical, operational, environmental and failure history.

Table 3-1 presents the influencing factors of pipe deterioration which should be collected for the application of the proposed methodology. Although these represent potential factors to assess the physical condition, some of these network properties are unknown for most water distribution networks. Therefore, it is recommended that, at least, the variables presented in 'bold' are collected, as these are considered as highly influential of pipe deterioration.

Table 3-1 – Potential factors influencing pipe deterioration.

Type of factor	Description
Physical	**Pipe material, year of installation, nominal diameter, length,** installation depth, wall thickness, Hagen Williams coefficient, cathodic protection, manufacturing process, location, laying conditions, number of households supplied.
Operational	**Static pressure,** pressure variance, water pH, chlorine levels, phosphate inhibitors, water temperature, water pressure, pressure surge allowance.
Environmental	Type of soil, ground water level, soil pH, soil resistivity, soil density, backfill material, traffic intensity, wheel load ratio, soil aeration index, road surface type, external loads, ratio of horizontal to vertical pressure in the soil.
Failure history	**Failure history**, date of burst, corrosion depth, bursting tensile strength, ring modulus of rupture.

Note: the collection of factors in bold are recommended for the application of the proposed methodology.

Before data analysis, the information gathered needs to be validated and checked for any inconsistencies (e.g., incorrect installation year of water pipes, inappropriate material choice for the diameter presented). These inconsistencies can be found through a detailed study of the network, in which the analysis of network properties versus historical and geographic data shows existing discrepancies. In most cases, these inconsistencies are the result of human error and can be corrected with the review of other network characteristics. A common example of inconsistencies found is the incorrect material associated to a pipe constructed in a given installation date, such as the using asbestos cement pipes in the last 20 years. The correction made to the data must reflect the appropriate nature of the assets that, without field survey, is carried out by analysing the characteristics of the surrounding pipes. Moreover, data validation also allows to audit existing records, which can contribute to the success of the overall infrastructure asset management program.

3.4 Step 2 – Model construction

The second step of the proposed methodology is the construction of assessment models of the physical condition of pipes. In this step, two mathematical modelling approaches are proposed: a deterministic and a heuristic approach. In both cases, an attempt is made to translate the collected factors into quantitative dependent variables (i.e., the failure rate in the deterministic approach and a non-dimensional condition value in the heuristic approach) that, in the final step, will allow the assessment of the physical condition of pipes. A detailed explanation of the deterministic and heuristic approaches is presented in the following sections.

3.4.1 Deterministic approach

The deterministic approach can be considered a data-driven approach that can be incrementally improved if new data are collected. This approach assumes that pipe deterioration has a non-random relationship with network characteristics that yield pipe failure. These failures can only be detected by water utilities during an inspection of pipes or when a burst occurs. The pipe burst is generally the most common form of failure history, as direct action needs to be taken when detected. As a result, this approach uses failure history, as dependent variables for model construction and aims to build models that estimate a theoretical failure variable for the installed pipes. The steps that are needed to complete this approach are the following:

i. The development of a correlation matrix between network characteristics and failure variables.
ii. The regression model construction.
iii. The validation of the regression models.

Firstly, to construct a correlation matrix, the used dependent and independent variables must be defined. As referred, the used independent variables are the characteristics of water distribution pipes (e.g., year of installation, material, nominal diameter) and the dependent variables are the failure history in the deterministic approach. Five dependent variables are defined as presented in Table 3-2. Preferably, the collected variables should be quantitative; this might require the conversion of categorical variables to a numerical value. An example of a categorical variable is pipe material, whose conversion to quantitative values can be carried out using the reference values of service lives for the different materials (see Table 2-5) or by other association of material properties with numerical values.

Table 3-2 – Dependent variables used for the proposed methodology.

Dependent variable	Description	Formula
Number of bursts (N_{Burst})	Total number of bursts observed	
Number of bursts per year *(NBY)*	Ratio of the number of bursts by the period of observation	$\left(N_{Burst}/PO\right)$
Average age at burst (*AA*)	Average age at which the pipe experienced a burst	
Rate of burst (*RoB*)	Rate of burts per meter*	$RoB = \frac{N_{Burst}}{l \times PO}$
Number of bursts in pipes (*AA03*)	Rate of burst per 100 km	$AA10 = \frac{N_{Burst}}{l \times PO} \times 100km$

Note: *l* is the length of the pipe in kilometres, PO is the observation time in years; *Defined by Alegre, Matos, *et al.*(2021)

The identification of the existing relationships between pipe characteristics and failure variables is carried out through the construction of a correlation matrix using the Pearson's product-moment-correlation-coefficient (PPMCC) or the Spearman's rank correlation coefficient (SRCC). PPMCC was proposed by Bravais, (1846) and measures the strength of the linear association between two variables (Laerd Statistics, 2020) and is calculated through the ratio of the covariance of the two variables by their product of the standard deviations, as follows (Morgan *et al.*, 1980):

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{3.3}$$

in which $\rho_{X,Y}$ is the PPMCC between the variables X and Y, $\mathrm{cov}(X,Y)$ is the covariance of the variables X and Y, σ_X is the standard deviation for the variable X and σ_Y is the standard deviation for the variable Y.

For both PPMCC and SRCC, correlation values range between -1.00 and +1.00, where 0 indicates no correlation between the variables, positive values indicate a direct linear relationship between the variables, which are stronger when close to +1.00, and negative correlation values indicate an inverse relationship between the variables (Morgan *et al.*, 1980). Although PPMCC is the most appropriate coefficient for the construction of the correlation matrix due to its linear relationship association, the variables used must be normally distributed. This should be considered since non-normally distributed variables increase the rate of Type I errors (false positives) and reduce robustness in a Pearson's correlation (Bishara & Hittner, 2012). Hence, the use of SRCC is the optimal alternative as it provides a nonparametric technique and measures the strength and direction of a non-monotonic relationship. In addition, the use of discreet variables, such as nominal diameter, makes SRCC a better alternative (King & Eckersley, 2019).

To calculate SRCC between two variables of size *n* (being *n* integer values), value pairs of the variables are converted to ranks ($R_i;S_i$) and evaluated using the following equation (King & Eckersley, 2019):

$$r_s = 1 - \frac{6\sum_{i=1}^{n} D_i^2}{n^3 - n} \qquad D_i = R_i - S_i \tag{3.4}$$

in which r_s is the Spearman rank coefficient.

The construction of the correlation matrix allows to identify the highest correlation coefficients that best describe the relationship between dependent and independent variables. This identification is essential to proceed to the second step of the proposed methodology, in which the models are constructed using the two different approaches. In this step, it is essential to confirm the goodness-of-fit of the models and their statistical significance. This analysis includes calculating the coefficient of determination, also known as r-squared, which evaluates the proportion of variance in the dependent variable that is predicted from the independent variable:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - f_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3.5}$$

in which n is the number of data points, y_i is the dependent value, f_i is the fitted value predicted by the value y_i and $\bar{y}$ is the mean of the data. The coefficients range between 0 and 1 and values approximate to 1 indicate a near-perfect fit, which represents a strong linear relationship between the variables, whereas values nearing 0 indicate a poor linear relationship between variables (Ramzai, 2020). Furthermore, the construction of multivariable linear models–requires the evaluation of the adjusted r-squared:

$$R^2_{adj} = 1 - (1 - R^2)\frac{n-1}{n-p-1} \tag{3.6}$$

in which R^2_{adj} is the adjusted r-squared, R^2 is the r-squared (equation 3.5), n is the number of data points and p is the number of dependents (or explanatory) variables. The usage of the adjusted r-squared aims to avoid the cumulative significance of independent variables and the values vary between 0 and 1, with the same meaning of r-squared.

In addition to the calculation of how well linear regressions adjust to the data points (r-squared), the statistical significance of the linear models is carried out by null-hypothesis testing. This is achieved by defining the significance level, α, and rejecting the null hypothesis if the p-value is lower than α. The *p*-value is calculated either by a t-test table or aided by software (J. Rumsey, 2016). The value of statistical significance defined for this study is 0.05 as it is a common standard in the urban water field.

Finally, the last statistical analysis is to verify independence between explanatory variables. The mathematical measurement used to perform this verification is the variance inflation factor (VIF), which quantifies the degree of multicollinearity in a regression analysis. The factor is calculated using equation (3.7), in which independent variables with values greater or equal to 10 are considered to have a high degree of multicollinearity and, therefore, should be excluded from the models (Neter *et al.*, 1983):

$$VIF_i = \frac{1}{1 - R_i^2} \tag{3.7}$$

in which VIF_i is the variance inflation factor for the variable *i* and R_i^2 is the R-squared of the variable *i* in the linear regression.

Before the third step of this approach, it should be evaluated which models present the most accurate results, since the application of the subsequent steps is unnecessary in models that do not meet the defined statistical criterion. Furthermore, if a larger number of constructed multiple linear models is obtained, the threshold values should be narrowed (i.e., decrease the *p*-value threshold for rejecting the null hypothesis, decreasing the acceptable degree of multicollinearity from ten to five), as a means to decrease the overall time of application of the methodology. Consequently, the narrowing of the statistical threshold values can improve the overall robustness of results. As described by Colquhoun (2014) the use of a *p*-value=0.05 for the rejection of the null hypothesis allows for a false discovery rate to be 30%, which can be avoided when using *p*-value ≤0.001, decreasing the false discovery rate to less than 5%.

The third and final step of the deterministic approach is the validation of the linear models. This involves verifying the statistical dispersion of the acquired data, namely through the calculation of the root mean square error, the relative error or the absolute error. To carry out this validation, the collected data must be divided in two subsets: the training subset, typically representing 80% of the whole dataset, which is used for model construction; and the validation subset, representing the remaining 20% of the dataset, which is used to validate the obtained results. The goal of the validation subset is to perform cross-validation of the evaluated models and flag problems, such as overfitting and selection bias (Cawley & Talbot, 2010).

3.4.2 Heuristic approach

A heuristic approach is by definition an approximate calculation method that aims to bypass strict nature of some types of algorithms, namely, stochastic methods, and achieve results (Dale, 2015). This approach does not necessarily produce better results than other methods. However, the main advantage is the ability to apply intuition to problem solving and to reduce the resources needed to achieve a satisfactory result (Foulds, 2014). The proposed heuristic modelling approach uses a theoretical framework as a basis for model construction. A survey of experts in the urban water field is carried out to determine the influence of internal and external factors on the deterioration of water distribution pipes. The key advantage of this modelling approach is that it is not limited to the observation period of the recorded failure history.

The steps to complete the heuristic approach are the following:

i. Survey development on the influence of internal and external factors on pipe deterioration.
ii. Processing and analysis of the survey data to calculate the weighting factors.
iii. Normalization of the main influencing factors associated with the pipe physical condition.

The first step aims to develop a questionnaire for experts to assess the relative importance that different factors have on the physical condition (translated by the deterioration) of water distribution pipes. The questionnaire should be short, clear and unambiguous and the structure chosen should facilitate data collection. Furthermore, it is relevant to ensure that experts only answer questions regarding factors they are familiar with, avoiding collection of guess-data. Additionally, it can be interesting to collect the age of surveyed experts or the number of years they have on the field to assess the perception that each experience group has on the influencing factors.

The developed questionnaire is presented in Appendix A. In this questionnaire, the influencing factors are evaluated on a scale from zero to five, in which zero means that the factor is irrelevant to pipe deterioration and five means that the factor is very relevant. The factors are grouped into categories, namely: (i) pipe characteristics and operating conditions; and (ii) external factors.

Following the collection of expert judgment, the next step is the processing and analysis of the acquired data to attain numerical weights that describe the influence that each factor has on physical condition. To achieve these weights, different statistical approaches can be taken to value different aspects of the collected data, namely the discrepancies in the mean value (i.e., the usage of the mean and median values) and the overall confidence of experts in their perceptions (i.e., the bias of the number of replies for each factor). The used weight methods are presented in Table 3-3. Furthermore, it should be noted that weights should range from zero to one and the total weight sum should be one.

The final step of the heuristic approach, it is necessary to normalize the variation of the acquired variables and of network properties for each pipe using the following equation:

$$N_{n,i} = \frac{y_{n,i} - y_{\min,i}}{y_{\max,i} - y_{\min,i}} \times W_i \tag{3.8}$$

in which $N_{n,i}$ is the normalized and weighted influencing factor i for pipe n; $y_{n,i}$ is the value of the influencing factor i for pipe n; $y_{\min,i}$ is the minimal value for the influencing factor i; $y_{\max,i}$ is the maximum value for the influencing factor i; W_i is the attributed weight for the influencing factor i.

Finally, the summation of the weighted variables of each factor results in a non-dimensional value of physical condition for each pipe as follows:

$$C_n = \sum_{i=1}^{n} N_{n,i} \tag{3.9}$$

in which $N_{n,i}$ is the normalized and weighted influencing factor i for pipe n and C_n is the condition value for the pipe n.

However, when applying equation (3.8), it should be noted that not all variables have the same relationship (positive or negative) with pipe physical condition. This is special attention should be paid to combine the variation of the influencing factor and their contribution to the condition grade. This is done by revisiting the knowledge gathered in section 2.5.2 regarding the factors that influence the deterioration of pipes and translating their deterioration effects in the normalization of the variables.

Table 3-3 – Methods for influencing factor weight attribution in the heuristic modelling approach.

Method	Equation	Description
Simple weight	$w_i = \frac{S_i}{\sum_{i=1}^{n} S_n}$	w_i : weight of factor i ; S_i : Sum of answers of factor i ; n : total number of influencing factors
Sum x Replies	$w_i = \frac{S_i \times A_i}{\sum_{i=1}^{n} S_i \times A_i}$	A_i : number of answers of factor i
Median	$w_i = \frac{m_i}{\sum_{i=1}^{n} m_i}$	m_i : median value of factor i
Median x Replies	$w_i = \frac{m_i \times A_i}{\sum_{i=1}^{n} m_i \times A_i}$	
Mean	$w_i = \frac{l_i}{\sum_{i=1}^{n} l_n}$	l_i : average value of factor i
Mean x Replies	$w_i = \frac{l_i \times A_i}{\sum_{i=1}^{n} l_i \times A_i}$	

3.5 Step 3 – Development of condition rating scale

The third step of the proposed methodology aims to develop a condition rating scale to interpret the results of the previously developed models and to assess the physical condition of water distribution pipes. A condition assessment matrix with three condition rating levels is developed: good, average and unsatisfactory. The condition matrices are developed based on the principles of risk matrices described by Cox(2008). For each approach (i.e., deterministic and heuristic) a different condition assessment matrix is developed with the modelled values being paired with the performance indicator IVI. The premise behind the matrices arises from the ambition to combine IVI with data related to observations (e.g., pipe burst). IVI is generally used for long-term investment planning, but can also be used to assess the physical condition of infrastructures considering the current value and their replacement cost. However, this performance indicator does not consider the good O&M practices and past interventions that have already been implemented. It corresponds to a well-known performance indicator that is often used by Portuguese water utilities and the water regulator (ERSAR), whose combination with the dependent variable evaluated in the deterministic and heuristic approaches allows a broader understanding of the network pipes' condition. Furthermore, it is a known fact that the calculation of IVI is unaffected by observations made during the network operation as it is the result of the ratio between the current value of the infrastructure and its replacement cost. The combination of IVI with a variable that translates a form of pipe performance during operation creates a more robust condition assessment method as it condenses different information sources into an overall assessment.

In both approaches, the calculated variable through the developed models is related to individual pipes. Consequently, the calculation of IVI must also follow the same rational. However, when calculating IVI at the asset level, the key performance indicator represents the ratio of useful life (RUL) of each pipe corresponding to the ratio between the asset age and the service life. Therefore, the RUL is incorporated in the condition matrices since the analysis is developed for each water distribution pipe.

The developed condition matrices for both approaches use the value of RUL in the y-y axis and adopts a division into three equal intervals of the variable. The RUL scale varies between zero and one, being the value one associated with newly installed pipes and the value zero associated with pipes that have already reach their reference service life and still remain in service. The original values of the IVI scale can be used (see Table 2-6), however, the condition assessment matrix would be unbalanced, as the average category is smaller than the remaining categories.

The x-x axis for each approach is the calculated variable through the previously developed models. Hence, the established condition intervals are different in each approach. In the heuristic approach, the same strategy for the RUL scale is proposed as the interval of variation of the evaluated variable is divided into three equal intervals. In the heuristic approach, the variation interval of the dependent variable varies between zero and one, however, the values at the extreme ends of the scale are very unlikely to occur as their occurrence implies an extreme scenario, in which an evaluated pipe is minimally affected by the weight of most of its evaluated property. In other words, the scenario where a single pipe is in most categories unaffected by the evaluated weight is very unlikely. This is in virtue of as some network properties work in opposite directions. For example, a pipe that has a high operating pressure is prone to be heavily affected by its weight but can be the youngest pipe in the network, causing it to be less effected by the weight of its year of installation. The proposed condition assessment matrix for the heuristic approach is presented in Figure 3-2(a).

In the case of the deterministic approach, the development of the x-x axis of the condition assessment matrix requires the prior evaluation of the acceptable condition thresholds considered for each water utility. This prior step requires the evaluation of the perception of pipe condition given the observation period. In the case of the chosen dependent variable being the number of bursts in the deterministic approach, the observation period is key to understanding the condition levels the evaluating utility considers acceptable. In this case, the issue that needs to be addressed is "*what is the acceptable number of bursts considered given the observation period?".*

In the case of an observation period of 5 years (as the considered case study of *Quinta do Lago*), the acceptable number of bursts for a pipe in good condition is considered to be zero, one burst for pipes in average condition and two bursts for pipes in unsatisfactory condition. Figure 3-2(b) presents the proposed condition assessment matrix for the deterministic approach for the case study of *Quinta do Lago.* In the case of an extensive observation period, the acceptable number of bursts might increase and might decrease in a limited observation period.

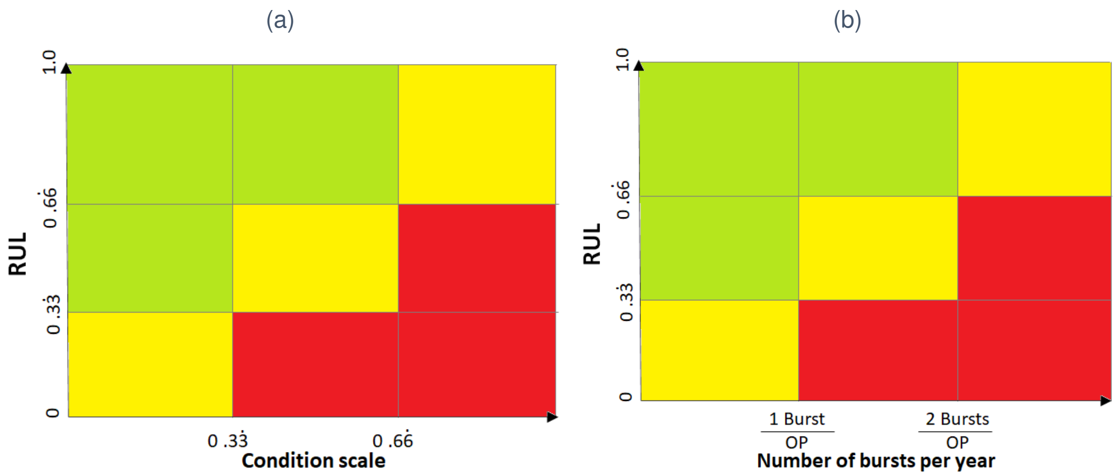


Figure 3-2 – Condition assessment matrices: (a) Heuristic approach; (b) Deterministic approach.

3.6 Step 4 – Condition assessment

The final step of the proposed methodology aims to assess the physical condition of the water distribution pipes. Firstly, the linear and polynomial models developed in the deterministic approach are applied to the entire network. The evaluated values of each model (number of bursts per year for the deterministic approach and the non-dimensional condition variable for the heuristic approach) are paired with their respective RUL value and plotted on the corresponding condition assessment matrix. This analysis allows to evaluate the physical condition of each pipe in the water distribution network using a variety of models and approaches that provide a basis for asset management. Furthermore, a colour grid map can be developed in a geographic information system, using the condition grades of each pipe colour coordinated with their respective condition to provide a visual aid for the overall network condition.

3.7 Summary and conclusions

This chapter presented the proposed methodology for assessing the physical condition of water distribution pipes using two different approaches: a deterministic and a heuristic approach.

The methodology consists of a four-step procedure, in which in the first step information regarding the characteristics of the pipe network is collected and validated, as well as information related to the failure history of water distribution pipes. This process of validation involves reviewing the collected data searching for inconsistencies, thus increasing the probability of correct modelling results in the following steps.

The second step describes the construction of models based on the deterministic and the heuristic approaches. The deterministic approach considers that the relationship between the pipe characteristics and failure history is a non-random process. Therefore, a failure history variable can be modelled, allowing the development of a condition rating scale that can be used to assess the physical condition of the pipes. On the other hand, the heuristic approach is theoretically based, aiming to develop a non-dimensional value of the physical condition for each pipe. This approach involves the survey of urban water experts on the influence of deterioration agents, referred to as influencing factors, on the deterioration of water distribution pipes. The collected data allows for the relative importance of each factor to be assessed through a weight, which describes the relationship between influencing factors and pipe deterioration.

The third step is the development of an appropriate condition rating scale for each approach. In the deterministic approach, the condition assessment matrix is constructed based on the evaluation of the appropriate service level acceptable by the water utility, while the condition assessment matrix for the heuristic approach is based on the division of the data range into equal intervals.

The fourth and final step assesses the physical condition of the water distribution pipes, applying the previous steps to the entire network. Results from Steps 2 and 3 culminate in this step, as the application of the developed models are necessary to classify the condition of the water pipes in each condition assessment matrix.

The proposed methodology can be applied to water distribution networks with different characteristics, allowing to assess the physical condition of buried pipes, without the need for inspection or extensive field survey. This methodology can consider a large number of different network characteristics. The following chapters present the application of the proposed methodology to the water distribution network of *Quinta do Lago.*

4 Case study characterization

4.1 Introduction

The current chapter describes the water distribution network of *Quinta do Lago*, the case study of this research work. A detailed description of the case study network including the characterization of the network and the failure history is carried out. The process of network characterization is an essential step since it allows to identify outliers in collected data.

4.2 InfraQuinta description

The water distribution network of *Quinta do Lago* is located in the south of Portugal, in Almancil municipality, in the district of Faro (Figure 4-1a). The network is located on a low-density residential housing (Figure 4-1b,c). The distribution network services 1787 houses with an annual volume of serviced water of 1 805 024 m^3. The distribution system includes three pumping stations (Infraquinta, 2019).

(a)

(b)

(c)

Figure 4-1 – Geography of case study location: (a) Map of Portugal with the case study location; (b) Overview of distribution network; (c) Landscape and urban development of the site.

The collected data regarding network characteristics and pipe failure history was provided by the water utility InfraQuinta, which is responsible for the maintenance and operation of the water distribution and drainage of this specific area (DR nº219, 2016). This water utility is a national reference due to its high-quality standards with assets in very good condition and leader in non-revenue water consumption in 2019 and 2020, with a percentage of total non-revenue water of

5.1% and 6.3% respectively (RASARP, 2020, 2021) The extensive data registry of network characteristics, as well as the use of an integrated geographical information system (GIS), allows the association of located bursts to the respective water distribution pipe. Consequently, this water utility was used for the application of the proposed methodology given the high infrastructure knowledge on their systems.

Additionally, the domestic consumers correspond, mainly, to detached houses (see Figure 4-1c), which are mostly used in the summer season. Thus, this utility is characterized by seasonal water consumption with high outdoor water uses in the summer and four to five times lower water consumption in the winter.

4.3 Network characteristics

The distribution network has a total length of 85 km and is composed of different materials, such as polyvinyl chloride (PVC), asbestos cement (AC), ductile iron (DI), steel and high-density polyethylene (HDPE). The most predominant pipe material is PVC, corresponding to 49% of the total network (Figure 4-3). The network is also composed of a large amount of asbestos concrete pipes (i.e., 35%), whose use was limited in 1987 (Limitação da Comercialização e Utilização do Amianto e dos Produtos que o Contenham, 1987) and prohibited in Portugal since 2005 (*Ministério Da Economia E Da Inovação*, 2005). The presence of asbestos concrete in water pipes does not cause a direct health risk to the serviced community if the material is in good physical condition (Direção-Geral de Saúde, 2020).

The nominal diameter varies between 60 and 600 mm and the maximum pipe length (as described in the GIS) corresponds to 4809 m. The length of each distribution pipe is obtained through a digital map, which may provide an approximate value of the pipe length. The analysed network is mainly composed of pipes with small nominal diameters (i.e., less than 250 mm), representing more than 75% of the total distribution pipes (Figure 4-2). This observation is unsurprising since the network services low-density residential housing.

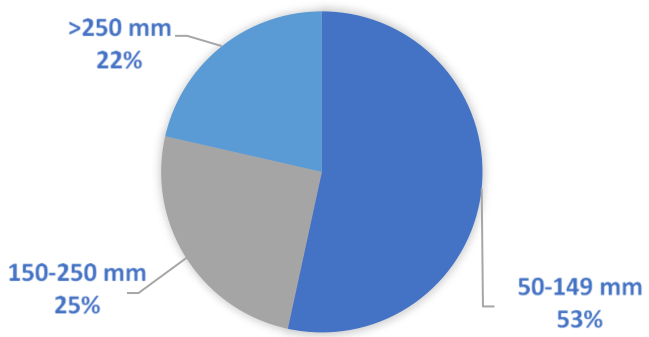


Figure 4-2 – Distribution of pipe nominal diameter in the distribution network.

The boxplot in Figure 4-3 describes the distribution of nominal diameters by pipe material, considering the pipes as defined in the GIS and not attending to their length. The ductile iron is the prominent material for very large diameter pipes, whereas the PVC are most used material in smaller diameter pipes. The number of HDPE and steel pipes is minimal.

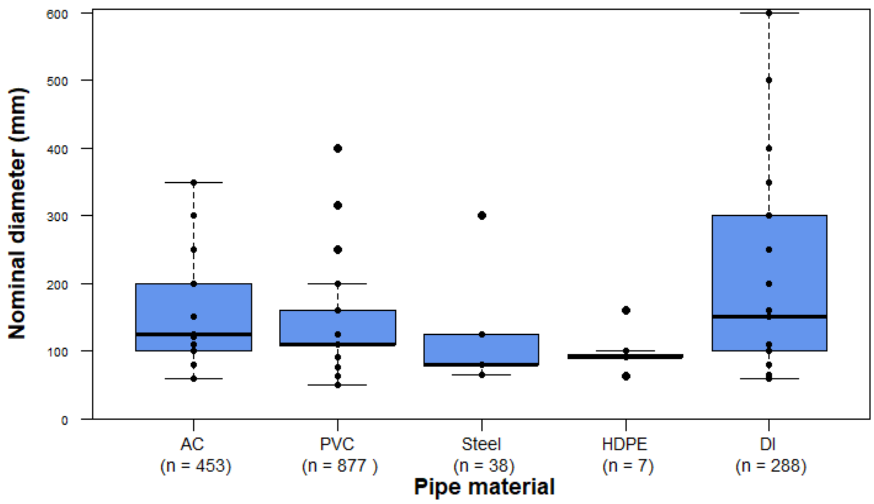


Figure 4-3 – Boxplot pipe material vs nominal diameter.

Only three PVC pipes with a total length of 1078 m and 46 DI pipes with a total length of 8887 m present a nominal diameter higher than 400 mm. The existence of these three PVC pipes can be justified by the advances in PVC manufacturing processes, which have increased the tensile strength and improved the durability of this material (British Plastics Federation, 2021). The possible reason behind the resistance to adopt PVC pipes in large diameters may be related to the fact that large diameter pipes are considered critical assets and, therefore, the concern for their durability and increased service life takes precedent that justifies their additional cost.

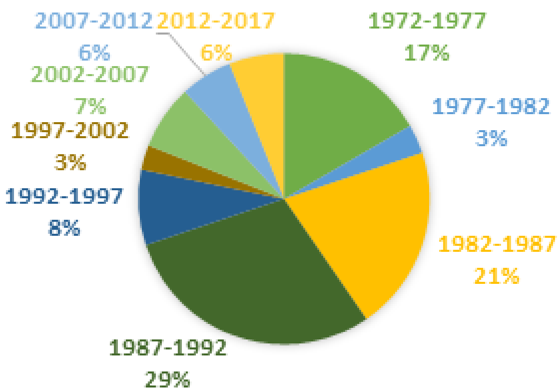


Figure 4-4 – Relative frequency of the installation year of the distribution network.

Regarding the pipes' construction year, the oldest pipes still in operation were installed in 1972, when the networks was initially constructed. The residential area was developed by André Jordan and designed by the architect João Caetano and the engineer Pedro Vasconcelos (Quinta do Lago, 2020). The development of the area was halted in 1975, only to be restarted in 1981. From 1982 to 1992, the largest expansion of the network took place, in which 50% of the pipes were installed, as depicts in Figure 4-4 . In recent years, no significant expansion of the network has been carried out, thus, the construction rate of new pipes is similar to the pipe replacement rate.

The materials used throughout the network are highly related to the pipe installation year. This relationship and can be associated with the evolution of the material choices for water distribution systems (Figure 4-5). The boxplots presented in Figure 4-5 show the distribution of the material type versus the installation year, ordered by increasing mean value. The points on each boxplot are one or more pipes of the designated material which were installed in the corresponding year, with *n* being the number of pipes of that material. Results have shown that AC, DI and PVC materials correspond to the oldest pipes. While PVC and DI are still installed, AC pipes have not been installed since 1993, as the manufacturing of asbestos products is linked to asbestosis, lung cancer and mesothelioma and, therefore, they fall into disuse (World Health Organization, 2000). Consequently, the materials of steel and HDPE became a choice for recent pipe replacement around 2010. Knowledge of the hazardous effects of asbestos was known since 1906 (Lean, 1996), but the high usage of this material persisted until the end of the 1980s and stops in the early 1990s. The decreased usage in the network can be linked to the entry of Portugal into the European Union in 1986 and the necessary compliance with EU laws by the early 1990s (Janela, 2017; European Union, 2020). Regardless, AC is still the second most common pipe material in the distribution network and, therefore, has to be addressed in condition assessment methods, even though it is no longer a material choice.

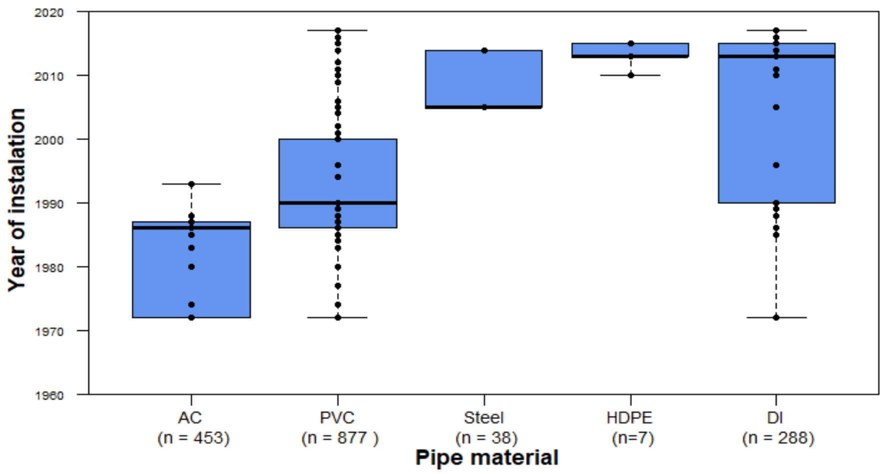


Figure 4-5 – Boxplot of pipe material by year of installation.

The percentage of pipe material installed per five years is described in Figure 4-6. The values presented corresponds to the total pipe length installed for each material. From the analysis of the material distribution in each five-year period, it can be concluded that the rate of network construction is not constant over time, with no relationship between material choice and network expansion. However, the shift from AC pipes to PVC is clear in Figure 4-6, as well as the shift to HDPE pipes in more recent years. Nonetheless, HDPE still remains one of the least common materials in the network. In the last decade, the HDPE material has been preferable for pipe replacement, thus, it is expected that AC pipes will be replaced by this material.

On the other hand, it can be concluded that AC was more abundantly installed in the years of higher expansion of the network (between 1982 and 1992). This conclusion can be explained by the fact that DI is the preferred material for large diameter pipes given its high reliability and structural resistance. Similarly, steel pipes are used for large diameter pipes, though their cost tends to be higher, and their installation requires more pipe fittings and specialized human resources for pipe welding. Thus, steel pipes are only used in the pumping stations located at downstream the water storage tank.

The collection and analysis of the network characteristics allow the validation of the data. The identification and correction of outliers in the network records was possible due to a deeper analysis of the network properties. Examples of the outliers found involved pipes with inappropriate nominal diameter to the associated pipe length, as well as the installation of AC pipes in the year 2005, which have been previously banned for commercial use. Upon the identification of inconsistencies in the data collected, the correction was achieved by consulting the surrounding pipes and validating the possible solutions with the water utility.

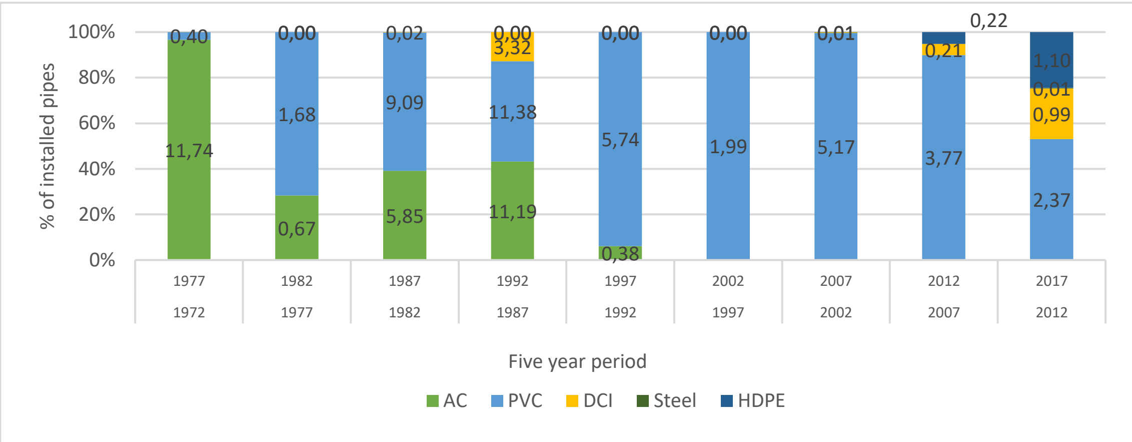


Figure 4-6 – Percentage of material installed per five-year period.

4.4 Failure history

The studied distribution network presented 113 bursts during the observation period, between February 2014 and February 2020. The water utility classified the registered bursts into eight different categories: excavation proximity, deficient construction or installation, active loss control, district metered areas (DMA) implementation, third party damage, burst, natural wear of materials and others. In this study, bursts related to third-party intervention or DMA implementation were disregarded. However, the bursts related to deficient construction or installation are a relevant category for this study as they can be associated with the burn-in phase of a water distribution network, as presented in the bathtub curve (see Figure 2-11).

From a total of 113 bursts, 41 were classified as third-party interventions, while the remaining 72 bursts can be attributed to the natural process of pipe deterioration. Two bursts were located in service connections and, therefore, were not considered in further analysis. Hence, only 70 bursts are relevant for the current investigation.

Similar to the network properties, the failure history also requires validation. However, the validation of the collected data requires access to the burst records of the water utility, which is unavailable for the purpose of this study. Therefore, the failure history data provided by InfraQuinta was presumed to be accurate, but future investigations should consider alternatives to this assumption.

Nevertheless, data analysis shows that a burst occurs, on average, once every 22 days. However, only one pipe presented five bursts during the observation period, being this the maximum number of bursts registered in a single pipe. The frequency of bursts in this pipe was between 68 days and 681 days. Some bursts can be aggregated into a single one, if they occurred within a 90-day interval (Jacobs & Karney, 1994). However, in the present case study, the aggregation of the bursts within a 90-day interval would decrease the number of available data points, which did not allow the application of the methodology.

A preliminary analysis of the network failure rates using the indicator AA03 (Table 3-2) described in Alegre, Matos, *et al.*(2021) is carried out, which classifies the network with a good quality of service, presenting 17 bursts per year per 100 km (out of a possible 30 bursts to achieve the same classification). However, the use of this indicator to assess the physical condition is limited to the assumption that the number of bursts per 100 km accurately describes the physical condition of the distribution network.

The number of expected bursts of a water distribution pipe tends to increase with its years in service. Table 4-1 supports this conclusion through the ratio of the number of bursts per total length. However, there are some exceptions. The high burst rate of AC pipes during the period of observation is likely due to the installation of this material mainly between 1972 and 1987. This is the result of the cumulative deterioration of the network pipes, which have exceeded their service life by 50%, according to reference service life presented in Alegre & Covas (2010)and exceed by 20% of their service life according to Covas, Cabral, *et al.*, (2018).

Table 4-1 – Number of pipe bursts per 5-year interval and associated ratio of the number of bursts per length.

Year of installation	Number of pipe bursts		Installed length [km]	Ratio of the number of bursts per length
	Absolute frequency	Relative frequency		
1972-1977	28	40%	12.14	2.31
1977-1982	4	6%	2.35	1.70
1982-1987	12	17%	14.97	0.80
1987-1992	17	24%	32.62	0.52
1992-1997	1	1%	6.12	0.16
1997-2002	2	3%	1.99	1.01
2002-2007	3	4%	5.19	0.58
2007-2012	3	4%	4.20	0.71
2012-2017	0	0%	4.47	0
TOTAL	70	100%	84.04	0.83

4.5 Summary and conclusions

The current chapter presented the characterization of the network of *Quinta do Lago* located in the Portuguese region of Algarve. The distribution network is associated with low-density residential housing with high seasonal water consumption. The network construction started in 1972, but only between 1982 and 1992 did most of the expansion take place. The most predominant materials in the network are AC, PVC and DI, as a result of being the most common materials during the expansion period.

A preliminary evaluation of the network failure rate using the AA03 indicator classified the network with a good quality of service. However, this indicator is limited to the overall assessment of the network and is not able to aid the assessment of individual assets. Furthermore, the failure rates of the network pipes present a strong relationship to the installation year and the pipe material.

5 Deterministic approach

5.1 Introduction

The current chapter aims at demonstrating the proposed methodology to assess the physical condition of water distribution pipes using the deterministic approach through the application to the water distribution network of *Quinta do Lago.* The chapter follows the steps of the proposed methodology for the deterministic approach including the considered assumptions, the data processing steps and the mathematical modelling.

5.2 Data collection and validation

The previous chapter characterized in detail the network of *Quinta do Lago* and, in this chapter, the assumptions and adaptations to the application of the proposed methodology to the deterministic approach are described. As referred, the deterministic approach aims to use network properties as independent variables to model the failure history of the network (e.g., the number of bursts, rate of bursts, number of bursts per year). The models are developed through the use of linear and polynomial regressions.

The observation period for the failure history of the studied network varies between February 2014 and February 2020, including a total of 70 bursts that were used for this research, being related to the natural deterioration process of distribution pipes. The bursts occurred in 49 different pipes which registered between one and five bursts. A summary of the dependent and independent variables used is described in Table 5-1.

Table 5-1 – Summary of the dependent and independent variables used in the deterministic approach.

Type of variable	Variable	Equation	Units	Abbreviation
Independent	Pipe material*		Years	**Mat5 or Mat3*** **
Independent	Installation year	-	Year	**YI or YI_0**
Independent	Length	-	m	**Len**
Independent	Static pressure	-	m	**SP**
Independent	Distance to the tank	-	m	**DT**
Independent	Nominal diameter	-	mm	**DN**
Independent/Dependent	Average age at burst	-	Years	**AA**
Dependent	Number of bursts	-	no.	**Nburst**
Dependent	Number of bursts per year	-	$\frac{Nburst}{year}$	**NBY**
Dependent	Rate of burst per meter**	$AA10 = \frac{N_{Burst}}{l \times PO} \times 100km$	$\frac{Nburst}{year \times 100km}$	**AA03**
Dependent	Rate of burst	$RoB = \frac{N_{Burst}}{l \times PO}$	$\frac{Nburst}{year \times meter}$	**ROB**
Dependent	Rate of burst length non-dimensional	$ROBLND = \frac{Nburst}{\bar{l} \times PO}$	$\frac{Nburst}{year \times meter}$	**ROBLND**

<u>Notes</u>: PO - Period of observation; l – pipe length;
*Converted to established service life;
**Defined by (Alegre *et al.*, 2021);

***Mat5 – Material 5 unique values Table 5-2 and Mat3 –Material 3 values Table 5-3

For both approaches (i.e., deterministic, and heuristic approaches), the collected pipe characteristics must be quantitative with a known relationship between them in order to develop the correlation matrix (Step 1) and the linear and polynomial models (Step 2). One of the independent variables that requires a transformation from a qualitative categorical variable to a quantitative variable is pipe material. This transformation is carried out based on established service lives of water distribution pipes (see Table 5-2). Although other associations can be made to develop this variable transformation, the decision to use the service life is based on the known relationship between the service life and the material durability. The agreed-upon service life for this study is presented in Table 5-2 and Table 5-3. However, since service lives are one of the most uncertain variables, additional service lives aggregated for the type of material (i.e., concrete, plastic or metal) are considered (Table 5-3). This aggregation is based on the known fact that some material types have similar service lives given their similar composition and durability.

Table 5-2 – Established service lives (considering singular values for each material).

Pipe material	Asbestos cement (AC)	Polyvinyl chloride (PVC)	High-density polyethylene (HDPE)	Steel	***Ductile iron*** (DI)
Established service life (years)	40	45	50	60	70

Table 5-3 – Established service lives (considering material similarities).

Pipe material	Asbestos cement (AC)	Polyvinyl chloride (PVC)	High-density polyethylene (HDPE)	Steel	***Ductile iron*** (DI)
Established service life (years)	40	50		60	

Furthermore, two variables associated with the year of installation were tested: year of installation (YI) and years in service (YI_0).

The variable distance to the tank (DT) is used as an indirect variable of pipe location, in order to consider the influencing factors attributed to environmental factors (e.g., traffic, soil pH, road surface type, external loads). No factors of this nature were provided for this research.

The variable rate of burst length non-dimensional (ROBLND) is used in this approach as an attempt to study the contribution of the length of each pipe to the rate of burst. This is a result of the observations made regarding the unbalanced range of pipe length presented throughout the work. The mean length of pipes is calculated to define ROBLND using Equation 3.1 to prevent the misrepresentation of the rate of burst caused by the defined pipe length in the datasets.

5.3 Model construction

5.3.1 Correlation matrix development

The construction of a correlation matrix aims to understand, at a preliminary stage, the relationship between the pipe characteristics and the failure variables. The use of Pearson's product moment correlation coefficient (PPMCC) requires that variables must be continuous, normally distributed, have a paired value and be absent of outliers that can skew relationships (King & Eckersley, 2019). Since some variables are discreet and only three of the 13 collected variables are normally distributed, the use of Spearman's rank order correlation coefficient (SRCC) is the most appropriate coefficient for the development of the correlation matrixes.

The construction of the correlation matrices only includes pipes that have bursts during the period of observation. Accordingly, the dataset is composed of only 49 elements (i.e., pipes) as opposed to 1666, which is the total number of pipes in the network. This reduced number of data points is preferable to avoid the excess noise created by pipes that did not experience bursts during the period of observation.

However, due to the low number of data points which can decrease the robustness of the modelling attempts, the decision to create two data sets was made. The first data set is composed of pipes that had bursts during the observation period (n = 49) and a second with both the pipes that experienced bursts during the observation period and the pipes that were installed during the observation period (n = 205). The choice to add the installed pipes to the second data set is associated with the certainty of pipe properties installed during this time.

Finally, before the construction of the correlation matrices, it should be pointed out that the data entry points are pipes and not single pipe bursts. Since the aim of the modelling approaches is to determine the physical condition of pipes and, therefore, the event of multiple bursts in a single pipe is indicative of poor physical condition. The correlation matrix for a data set of n = 49 is presented in Table 5-4 and the correlation matrix for the largest data set of n = 205 is presented in Table 5-5. The colour gradient described in these tables is increasingly red for inverse correlation and increasingly blue for direct correlation values.

The analysis of both correlation matrix demonstrates that there is a strong direct correlation between year of installation and material type. This conclusion is unsurprising given that material choices are associated with the evolution of pipe material. A high correlation between the variables of distance to the tank and static pressure is observed, which can be attributed to the natural geography of the network, since the distribution tank is located on a hilltop and the network progresses towards the sea-level, therefore, assimilating the static pressure in network pipes and the distance to the tank. Other network properties did not present such a strong relationship.

Table 5-4 – Correlation matrix using spearman's rank level correlation coefficient for Quinta do Lago – Only pipes with bursts (n=49).

	Mat5	Mat3	YI	YI_0	DN	length	AV_T	DT	SP	NB	AA03	AA	ROB	ROBLND	NBY
Mat5	1.00	1.00	0.60	-0.60	-0.13	-0.20	0.02	0.07	0.11	-0.38	0.06	-0.41	0.06	-0.38	-0.38
Mat3	1.00	1.00	0.60	-0.60	-0.13	-0.20	0.02	0.07	0.11	-0.38	0.06	-0.41	0.06	-0.38	-0.38
YI	0.60	0.60	1.00	-1.00	0.19	-0.13	-0.11	0.28	0.30	-0.51	-0.09	-0.45	-0.09	-0.51	-0.51
YI_0	-0.60	-0.60	-1.00	1.00	-0.19	0.13	0.11	-0.28	-0.30	0.51	0.09	0.45	0.09	0.51	0.51
DN	-0.13	-0.13	0.19	-0.19	1.00	0.18	-0.18	0.03	-0.04	0.12	-0.17	0.14	-0.17	0.12	0.12
Length	-0.20	-0.20	-0.13	0.13	0.18	1.00	0.03	0.13	-0.04	0.34	-0.91	0.09	-0.91	0.34	0.34
AV_T	0.02	0.02	-0.11	0.11	-0.18	0.03	1.00	0.02	-0.01	0.06	-0.05	-0.11	-0.05	0.06	0.06
DT	0.07	0.07	0.28	-0.28	0.03	0.13	0.02	1.00	0.84	-0.22	-0.25	-0.25	-0.25	-0.22	-0.22
SP	0.11	0.11	0.30	-0.30	-0.04	-0.04	-0.01	0.84	1.00	-0.35	-0.13	-0.38	-0.13	-0.35	-0.35
NB	-0.38	-0.38	-0.51	0.51	0.12	0.34	0.06	-0.22	-0.35	1.00	0.05	0.63	0.05	1.00	1.00
AA03	0.06	0.06	-0.09	0.09	-0.17	-0.91	-0.05	-0.25	-0.13	0.05	1.00	0.14	1.00	0.05	0.05
AA	-0.41	-0.41	-0.45	0.45	0.14	0.09	-0.11	-0.25	-0.38	0.63	0.14	1.00	0.14	0.63	0.63
ROB	0.06	0.06	-0.09	0.09	-0.17	-0.91	-0.05	-0.25	-0.13	0.05	1.00	0.14	1.00	0.05	0.05
ROBLND	-0.38	-0.38	-0.51	0.51	0.12	0.34	0.06	-0.22	-0.35	1.00	0.05	0.63	0.05	1.00	1.00
NBY	-0.38	-0.38	-0.51	0.51	0.12	0.34	0.06	-0.22	-0.35	1.00	0.05	0.63	0.05	1.00	1.00
See Table 5-1 for legend															

Table 5-5 – Correlation matrix using spearman's rank level correlation coefficient for Quinta do Lago – Pipes with bursts + Pipes installed during observation period (n=205).

	Mat5	Mat3	YI	YI_0	DN	length	AV_T	DT	SP	NB	AA03	AA	ROB	ROBLND	NBY
Mat5	1.00	0.98	0.66	-0.66	-0.14	-0.63	-0.80	-0.09	-0.02	-0.81	-0.79	-0.81	-0.79	-0.81	-0.81
Mat3	0.98	1.00	0.67	-0.67	-0.18	-0.61	-0.82	-0.11	-0.02	-0.83	-0.81	-0.83	-0.81	-0.83	-0.83
YI	0.66	0.67	1.00	-1.00	-0.20	-0.32	-0.75	0.06	0.10	-0.76	-0.75	-0.76	-0.75	-0.76	-0.76
YI_0	-0.66	-0.67	-1.00	1.00	0.20	0.32	0.75	-0.06	-0.10	0.76	0.75	0.76	0.75	0.76	0.76
DN	-0.14	-0.18	-0.20	0.20	1.00	0.16	0.06	0.03	0.03	0.08	0.06	0.08	0.06	0.08	0.08
Length	-0.63	-0.61	-0.32	0.32	0.16	1.00	0.62	0.06	-0.04	0.62	0.58	0.62	0.58	0.62	0.62
AV_T	-0.80	-0.82	-0.75	0.75	0.06	0.62	1.00	0.01	-0.07	0.98	0.98	0.97	0.98	0.98	0.98
DT	-0.09	-0.11	0.06	-0.06	0.03	0.06	0.01	1.00	0.82	0.01	0.02	0.03	0.02	0.01	0.01
SP	-0.02	-0.02	0.10	-0.10	0.03	-0.04	-0.07	0.82	1.00	-0.09	-0.05	-0.05	-0.05	-0.09	-0.09
NB	-0.81	-0.83	-0.76	0.76	0.08	0.62	0.98	0.01	-0.09	1.00	0.98	0.99	0.98	1.00	1.00
AA03	-0.79	-0.81	-0.75	0.75	0.06	0.58	0.98	0.02	-0.05	0.98	1.00	0.98	1.00	0.98	0.98
AA	-0.81	-0.83	-0.76	0.76	0.08	0.62	0.97	0.03	-0.05	0.99	0.98	1.00	0.98	0.99	0.99
ROB	-0.79	-0.81	-0.75	0.75	0.06	0.58	0.98	0.02	-0.05	0.98	1.00	0.98	1.00	0.98	0.98
ROBLND	-0.81	-0.83	-0.76	0.76	0.08	0.62	0.98	0.01	-0.09	1.00	0.98	0.99	0.98	1.00	1.00
NBY	-0.81	-0.83	-0.76	0.76	0.08	0.62	0.98	0.01	-0.09	1.00	0.98	0.99	0.98	1.00	1.00
See Table 5-1 for legend															

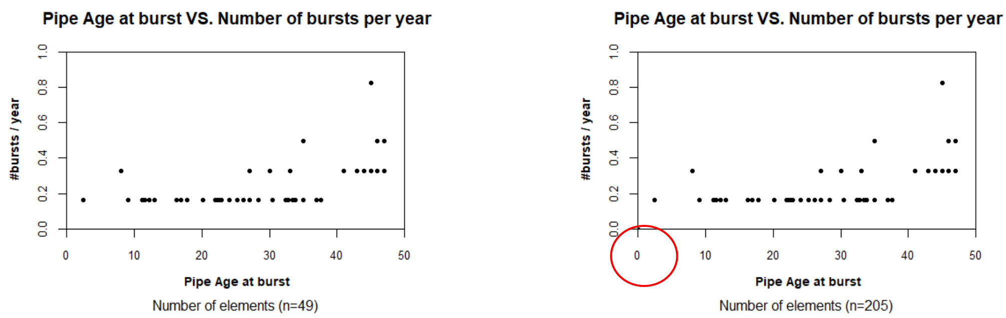
Furthermore, the analysis of the correlation matrices demonstrated the existence of two groups of failure variables: Group 1 (NB, NBY and ROBLND) and Group 2 (ROB and AA03). This conclusion is not surprising given that the difference between the grouped variables is constant values. In the first group, the number of bursts is divided by two constants (year and length), while in the second group the number of bursts is divided by different values of length. Upon this conclusion, the analysis of the relationships between the variables derived from the network characterization (i.e., year of installation, material, DN, length and static pressure) with the factors that can only be found in the failure history of the network (i.e., distance to the tank, average water temperature at the month of burst) can begin.

Finally, the correlation matrix obtained for the dataset n=49 (Table 5-4) shows that the reduced number of data points is sufficient to establish some relationships between the independent and the dependent variables, namely between failure variables and pipe material (Mat5), year of installation (YI), pipe length (length), distance to reservoir (DT), static pressure (SP) and average age (AA). However, the correlation matrix of the larger data set (n=205) presents even stronger relationships between the dependent and the independent variables. However, the dependent variables of AA03 and ROB present disproportionality large correlation value, which leads to suspect the large data set distorts the results and that the addition of extra pipes did not improve the insight into the physical condition of assets but rather cause the misrepresentation of the correlations. Furthermore, analysis of the correlation matrices demonstrates that the use of five or three material categories did not present different correlation results.

5.3.2 Construction of linear and quadratic models

Three possible dependent variables can be chosen for the development of the deterministic models: these are NB, NBY and ROBLND. The analysis of these three variables shows that the models constructed will be similar, as the only difference between them are constant values. Therefore, the decision to choose a single dependent variable out of these three no longer becomes statistically-based, but a technical choice. Therefore, the variable allows the clearest interpretation by the water utilities is the number of bursts per year (NBY), as such, this is chosen to be the dependent variable.

Following the choice of the number of bursts per year as the dependent variable for the present case study, the scatter plot of the real number of bursts versus the ratio of useful life is presented in Figure 5-1 for both datasets. The obtained plots for both datasets are practically identical. The only difference between them is illustrated in Figure 5-1b (see red circle), where 156 data points are overlapped in (0,0), as these are the pipes that did not experience bursts during the observation period. A preliminary analysis of the modelling attempts for the larger dataset presented elevated r-squared values. However, the reason behind the improved value is in virtue of the 156 points presenting small variance between them, causing an artificial increase in the coefficient of determination as the point (0,0) holds a strong influence over the model. Therefore, the increased significance in the results when using the data set n=205 does not represent an effective (real) improvement of the modelling results. Therefore, the previous suspicion to abandon the larger dataset is now supported by these results and the dataset of n=205 is abandoned from further modelling.



(a) (b)

Figure 5-1 – Scatter plots; (a) Pipe age at burst vs. number of bursts per year (n=49); (b) Pipe age at burst vs. number of bursts per year (n=205).

A thorough process of construction of the linear models involves testing all the combinations of explanatory dependent variable. The results for the r-squared values and adjusted r-squared values for all the possible linear models constructed from the chosen explanatory variables using the dataset n=49 is presented in Appendix B. The analysis of the linear models developed allowed to narrow the obtained models to the six best goodness-of-fit models presented in Table 5-6.

Table 5-6 – Results of the models with the highest coefficient of determination.

Model variables	R^2	R^2_{adj}	Y-intercept	β_1	β_2	β_3	β_4
AA+SP	0.51	0.48	0.33760	0.00717	-0.00565	-	-
AA + SP + Length	0.52	0.48	0.35870	0.00733	-0.00599	-0.00003	-
AA + Mat3 +SP	0.44	0.48	0.40384	0.00688	-0.00132	-0.00563	-
AA+YI+SP	0.51	0.48	1.52630	0.00688	-0.00060	-0.00547	-
AA+DTT+SP	0.52	0.47	0.35670	0.00716	0.00000	-0.00623	-
AA+Mat3 + SP + Length	0.44	0.47	0.45990	0.00698	-0.00186	-0.00609	-0.00004

AA – average age of pipe at burst; SP- Static pressure; Mat3 – Material type in 3 values; Length- Pipe length; YI – Year of installation; DT – distance to tank

Table 5-6 shows that the variable AA (average age of pipe burst) is present in all models and that this variable presents the highest correlation coefficient. In fact, when ordering the constructed models by the correlation coefficient (adjusted or not), from largest to smallest, the first 32 models use AA as a dependent variable. Thus, AA can be considered a variable with a high explanatory ability. Figure 5-2 presents the best fit linear model developed, where the lines represent a set of constant static pressure-heads.

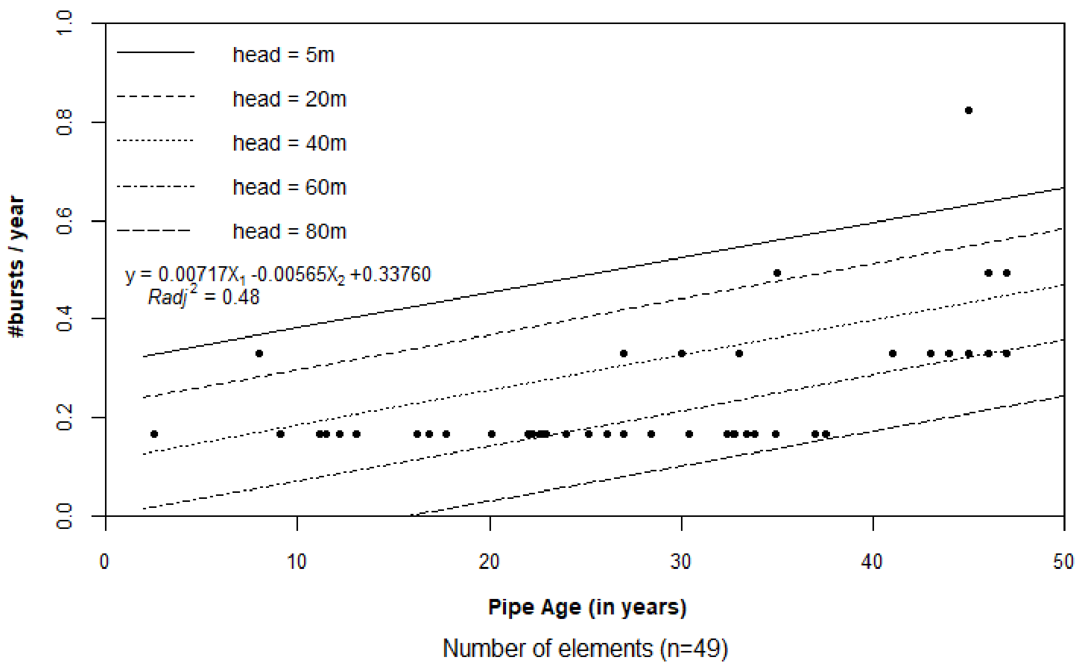
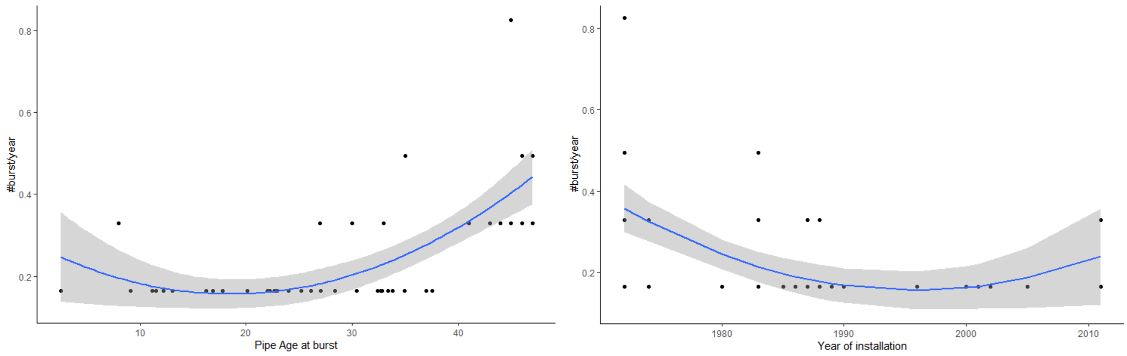


Figure 5-2 – Linear model of average age at burst and static pressure (AA+SP).

While the correlation values for linear models seem appropriate, the possibility of existing a quadratic function that better fits the number of bursts per year is explored. As the variables average age at burst and year of installation present promising correlation results for linear models, these variables are tested for polynomial fitting and the corresponding results are presented in Table 5-7. When analysing the obtained results versus the linear regression models, it can be concluded that no significant improvement is attained in this modelling attempt. The plotting of the modelled functions is presented in Figure 5-3 with possible intervals of variance in grey.

Table 5-7 – Equation and statistical results of quadratic equations for Average age at burst and year of installation.

X-Variables	R^2	Y-intercept	β	$β^2$
AA	0.50	0.28	-0.01	0.000352
YI	0.33	1419.00	-1.42	0.000356
AA – average age of pipe at burst; YI – Year of installation;				



(a) (b)

Figure 5-3 – Quadratic modelling: (a) Number of bursts vs age of pipe at burst; (b) Number of bursts per year vs. year of installation.

5.3.3 Variable independence

Following model construction, it is important to verify the statistical dependencies of the influencing factors. The variance inflation factor (VIF) is used to assess the degree of multicollinearity between the variables. The VIF value of 10 is considered a high degree of multicollinearity, consequently, variables that present approximate values to 10 are discarded. In addition, VIF value is determined not in relation with each variable, but in relation to the modelled dependent variable, therefore linear models for all the dependent variables were developed.

The results of the VIF analysis are presented in Table 5-8, for the first data set (n=49). The rows in Table 5-8 present the failure history variables that can be dependent variables and the columns present the independent variables. The variable average age at burst (AA) is used for both independent and dependent variable, but never simultaneously. The reason is that when it is used as dependent variable it intendeds to estimate the age at which pipes will burst, while, when applied as an independent variable is becomes the age at which the pipe has experienced bursts. The VIF analysis presented in Table 5-8 shows that all variables can be used to model the relationship between the pipes properties and failure history.

Table 5-8 – Variance influence factor analysis for data set n=49.

	Mat	YI	DN	Length	DT	SP	AV_Temp	AA	NB	ROB
AA	2.33	2.31	1.30	1.69	4.21	5.49	1.03	D	-	-
NB	2.33	2.93	1.38	1.75	4.28	5.91	1.07	1.74	D	-
ROBLND	2.33	2.93	1.38	1.75	4.28	5.91	1.07	1.74	A	-
NBY	2.33	2.93	1.38	1.75	4.28	5.91	1.07	1.74	A	-
ROB	2.33	2.93	1.38	1.75	4.28	5.91	1.07	1.74	-	D
AA03	2.33	2.93	1.38	1.75	4.28	5.91	1.07	1.74	-	A

Mat- pipe material; YI – year of installation; DN -pipe nominal diameter; SP- static pressure; AV_Temp – average temperature at burst AA- average age at burst; NB- number of bursts; ROB- rate of burst; NBY- number of bursts per year; AA03- rate of bust per 100 km; D - analysed variable; A- perfectly colinear variable.

5.3.4 Validation of dependent variable

The second step in the deterministic approach requires the selection of the most appropriate dependent variable to be modelled. The attempt to model all possible combinations of dependent variables follows the formula: 2^n for each dependent variable, where *n* is the number of combinatory variables. Therefore, the attempt to model the seven explanatory variables collected with the six possible dependent variables would result in 768 linear models that would unnecessarily increase the time required for the application of the methodology. Therefore, based on the correlation matrix results, the number of bursts per year as the dependent variable is chosen as the dependent variable. However, this decision requires further validation.

To validate the choice of dependent variable, bidirectional elimination in the form of a stepwise algorithm is performed to identify the best combination of independent variables for the three dependent variables in question (NB, NBY and ROBLND). The performed statistical test is the evaluation of the r-squared value and p-value, where the defined level of significance is 0.05, in accordance with the section 0 . This means that the model coefficients of R^2, adjusted R^2 and the p-value are assessed and considered deciding criteria for model exclusion. Table 5-9 presents the results of a stepwise analysis, in which the r-squared (R^2) and adjusted r-squared (R^2_{adj}) of a modelling attempt with all the variables are presented.

Table 5-9 – Stepwise analysis results n=49.

Observations	Dependent variable	Independent variable	R^2	R^2_{adj}
	NB	ALL	0.52	0.43
Stepwise		SP+ AA	0.51	0.48
	NBY	ALL	0.52	0.43
Stepwise		SP+ AA	0.51	0.48
	ROBLND	ALL	0.52	0.43
Stepwise		SP+ AA	0.51	0.48
	AA	ALL	0.41	0.30
Stepwise		YI+ DN+ SP	0.38	0.34
AA- average age of pipe at burst; SP- Static pressure; Length- Pipe length; YI – Year of installation;				

From the previous exclusion of the ROB and AA03 from dependent variable, as a result of poor correlation with pipe properties, the remaining possible variables to be used in the deterministic modelling attempt are NB, NBY, AA, and ROBLND.

Finally, Table 5-9 shows that the results obtained for NB, NBY, and ROBLAND are very similar. Therefore, the choice for a NBY as a consequence of its clearer interpretation of results by water utilities is supported and the assumption of using either one of the three variables is confirmed.

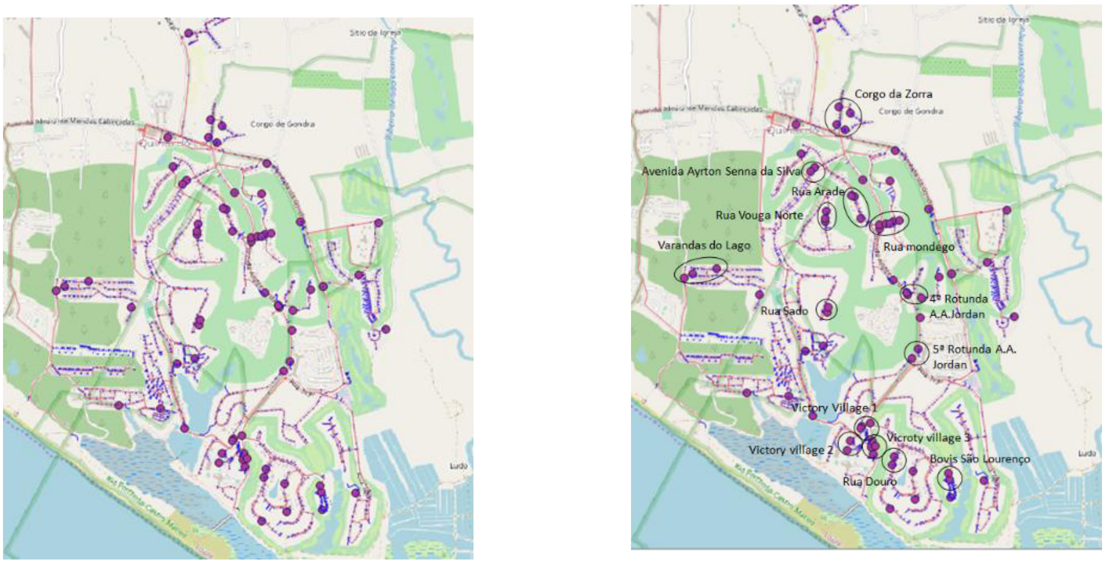
5.3.5 Clustering attempts

While mapping the correspondence between the location of bursts and the corresponding pipes during the first step of this methodology (5.2), it is possible to observe a tendency for clustering of pipe breaks in certain locations. Upon further analysis, it bursts are clustered in certain roads and intersections as seen in Figure 5-4a.

In the literature review, many authors proposed aggregating pipes into cohorts with similar properties as a way to increase statistical significance (Royer, 2012; Thomson & Royer, 2013; Ugarelli & Bruaset, 2010). As a result, it was hypothesized that cohorts could be formed based on spatial similarities to possibly account for the influence of environmental factors. Following this observation, the map in Figure 5-4(b) is created, aggregating pipe bursts into clusters. The rules for cluster construction are:

- pipe bursts are located within 150 m of a cluster centroid;
- pipe bursts belong to the same street of the network.

The application of these rules result in 12 different clusters that are referred according to the road or cul-de-sac they were associated with.



(a) (b)

Figure 5-4 – Maps of distribution network: (a) Map of distribution pipe with failure location; (b) Map of distribution network with aggregated clustered.

Following this aggregation into clusters, the properties of each pipe are compared to each other. This comparison reveals that the clusters are fairly homogenous in terms of pipe material and diameter. This conclusion allows the clusters to be defined by their most dominant attributes and added as new variables to the existing correlation matrix. As a result, each data point has now three new properties regarding the distance of the cluster centroid to the tank, the cluster material and the number of bursts per cluster. The results obtained through the Spearman's rank correlation coefficient are presented in Appendix B - Construction of linear models ().

After analysing the results, the correlations between pipe burst and distance to tank prove to be significant within clusters. However, there is no significant improvement in correlation from the clustering attempt, which justifies the aggregation. Therefore, the possibility of pipe aggregation into cohorts is not considered for this case study and, therefore, not part of the proposed methodology.

5.4 Development of condition rating scale

Following the construction of relevant deterministic models, it is now appropriate to define the condition thresholds through a condition assessment matrix. As described in section 3.5, to do so, it is necessary to recall the knowledge gathered concerning the infrastructure value index (IVI) and the condition thresholds associated with that metric (see

Table 2-6). IVI is a value-based index, defined by as the ratio between the current value of the infrastructure and the replacement cost (see equation 2.2). This indicator was originally proposed to be used as a metric for long-term planning (Alegre *et al.*, 2014) as well as to provide a general perspective of the condition of assets (Cabrera Rochera *et al.*, 2019), However, if the IVI calculation is used at an individual asset level, the index resembles the ratio of useful life (RUL) which identifies the percentage of remaining useful life of an asset. Consequently, for the development of the condition assessment matrix, this term is selected with the addition of some the underlying assumptions of IVI.

The direct adoption of the condition intervals defined by IVI (see Table 2-6) for the development of the condition assessment matrix cause an imbalanced matrix with a narrow interval for pipes with average condition which would result in the overestimation of deterioration rates of certain pipes. For this reason, more evenly distributed intervals are adopted in this analysis, as presented in Table 5-10.

Table 5-10 – Condition intervals for ratio of useful life (RUL) values.

Condition	RUL value range
Good	[1.00 – 0.66[
Average	[0.66 – 0.33]
Unsatisfactory	]0.33 – 0.00]

The development of the horizontal axis of the condition assessment matrix requires the prior evaluation of what is considered by the water utility to be good, average and unsatisfactory condition given the chosen dependent variable. Thus, a quantification of each condition level is carried out, which can be revaluated with each application of the proposed methodology.

As referred, the chosen dependent variable is the number of bursts per year, therefore, the prior evaluation required for this application of the methodology dictates that the number of bursts per year must be interpreted to allocate each condition level. The number of bursts per year that are found to be acceptable relied on the broader evaluation of the number of bursts during the 6-year observation period. For this observation period, it is assumed that i) pipes with more than two bursts are considered to be in unsatisfactory condition; ii) pipes with one or two burst can be classified as average condition; and iii) pipes with no bursts during this period are considered to be in good condition. The application of these assumptions resulted in condition intervals presented in Table 5-11.

Table 5-11 – Proposed condition levels for the number of bursts per year.

Condition	Number of bursts in 6 years	Number of bursts per year
Good	0	[0.00 – 0.15[
Average	1-2	[0.15 – 0.30]
Unsatisfactory	> 2	]0.30 – 0.45]

The combination of the values presented in Table 5-10 and Table 5-11 results in the condition assessment matrix depicted in Figure 5-5. This matrix also shows the scatter plot of the RUL values versus the number of bursts per year from the case study. Pipes with a low RUL and a high number of bursts per year are clearly classified as unsatisfactory condition (red area). The use of the condition assessment matrix grants the possibility of pipes that are at extreme ends of each axis to be given the benefit of the doubt regarding their condition. An example of this is a pipe with two bursts and a RUL higher than more than 0.6 that can be categorized in average condition, as opposed to being unsatisfactory condition if only the former criterion is used.

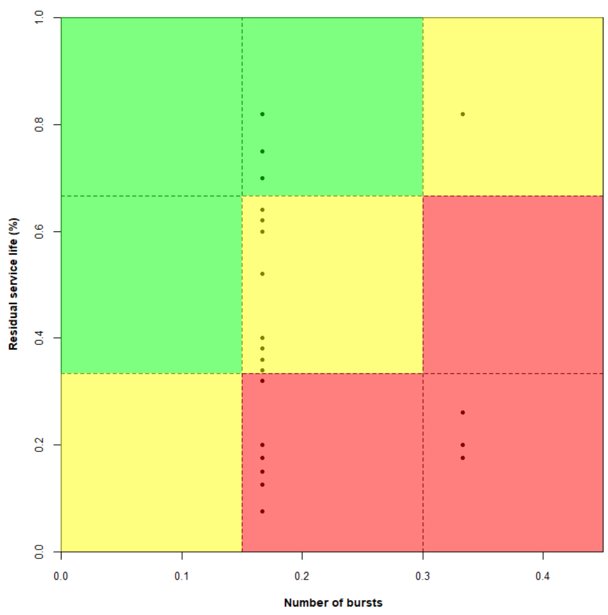


Figure 5-5 – Condition assessment matrix for collected values

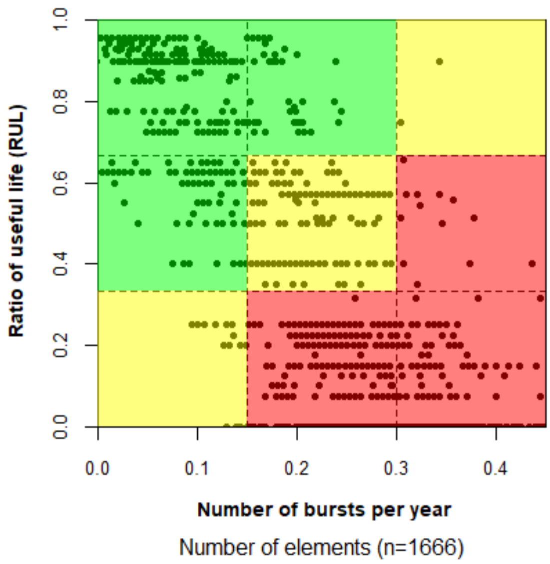
5.5 Condition assessment – Deterministic approach

The fourth and final step of the proposed methodology is the condition assessment of each water distribution pipe using the developed models and the condition assessment matrix developed in the previous step. When applying the developed models to the entire network pipes, the variable average age at burst (AA) only has exists for pipes with bursts in the observation period. Thus, the dependent variable years in service instead is used instead. Yet, this decision eliminates one of the six best deterministic models, as it used both AA and year of installation which made the model redundant.

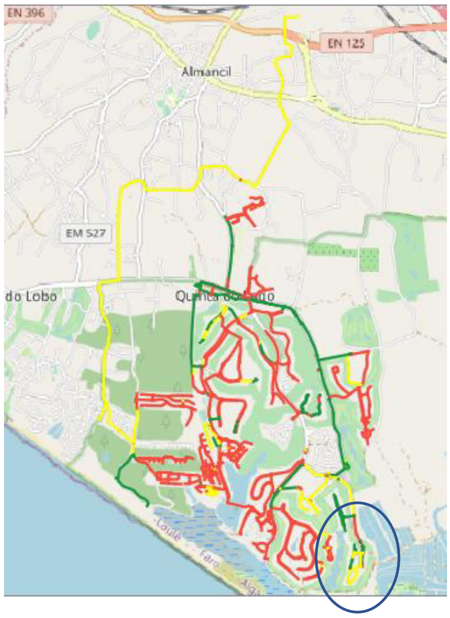
Results of the deterministic approach using the a linear model with the variables average age at burst and static pressure (AA+SP) are presented in Figure 5-6: Figure 5-6(a) presents the results for the evaluation of each pipe using the developed condition assessment matrix and Figure 5-6(b) presents a map of the network with each pipe colour coded for its condition level. Figure 5-7 presents the same results for the deterministic approach but using the polynomial model of average age at burst (AA^2). Table 5-12 presents the summary of the pipes classified in each category by the two deterministic models.

The comparison of the results obtained by the two models shows that the linear model classifies more pipes with good condition than the polynomial model. However, both models classify almost 60% of the network with an unsatisfactory condition. This conclusion is surprising when considering that the network of Quinta do Lago is considered to be in good condition when applying the indicator AA03 (Alegre *et al.*, 2021).

Appendix C presents the modelling results for the remaining linear and polynomial models. The comparison between the results from using the quadratic and the linear models are overall quite similar, in particular in unsatisfactory condition, as both models tend to overestimate the level of deterioration the water distribution pipes. However, the pipe classifications show small differences in the other two categories. An example is the analysis of the pipes located in the bottom right-hand side in Figure 5-7(b) and in Figure 5-6(b) (see detail). When observing the colour grid map, the quadratic model classifies that area with an average condition grade, while the linear model is much more diverse in its classification.

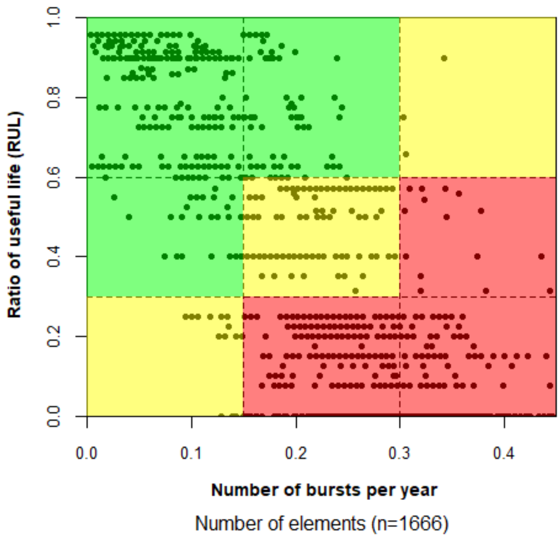


(a)

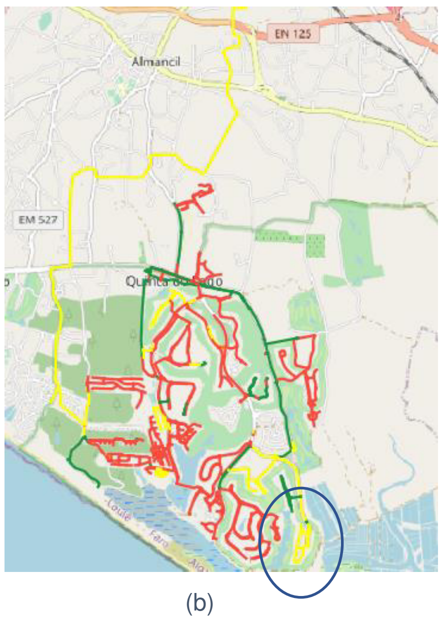


(b)

Figure 5-6 – Results of the deterministic linear model AA+SP: (a) pipe condition based on the condition assessment matrix; (b) map of network pipes. Detail: blue circle.



(a)



(b)

Figure 5-7 – Results of deterministic quadratic model AA^2: (a) pipe condition based on the condition assessment matrix; (b) map of network pipes. Detail: blue circle.

Table 5-12 – Deterministic results for model summary: Average age at burst + Static pressure.

Condition grade	**Linear model (AA+SP)**		**Quadratic model (AA^2)**	
	Number of pipes	**(%)**	**Number of pipes**	**(%)**
Good	476	28.57%	342	20.53%
Median	194	11.64%	346	20.77%
Unsatisfactory	993	59.60%	978	58.70%
Total	**1666**	**100%**	**1666**	**100%**

5.6 Summary and conclusions

The deterministic approach is a highly intensive data-driven approach that yields results mostly based on the provided data. The primary advantage of this approach is the minimization of the subjective nature of condition assessment. The process of data collection and validation can be labour intensive and do not guarantee accurate results. The presence of outliers that are not removed can have a negative impact on the model results and distort the conclusions.

Finally, the overall results of the deterministic approach show that this method may overestimate the estimated deterioration level of the water distribution pipes. The hypothesis behind this statement is that the developed models are based on a dataset that only includes pipes with bursts and, when considering that pipe bursts are a physical manifestation of the unsatisfactory physical condition of water distribution pipes, this deterministic assumes that all other pipes will deteriorate in the same manner as those that presented the worst physical condition. Consequently, these assumptions present the worst-case scenario that is difficult to overcome given that the baseline for model construction is the pipes that had bursts (i.e., those with the poorest condition).

6 Heuristic approach

6.1 Introduction

The current chapter presents the application of a heuristic approach to the case study of *Quinta do Lago.* The chapter follows the steps described for the heuristic approach (section 3.4.2), which is based on the knowledge of experts in the field and on several methods for the weights calculation. A survey with more factors than those collected in the *Quinta do Lago* case study is carried out in order to understand the perception of experts on a large number of influencing factors.

6.2 Data collection and validation

Data collection and validation includes the evaluation of the current knowledge in the urban water field regarding the influence of several factors on asset deterioration. While the deterministic approach is based on the evidence provided by the collected data, the construction of models using a heuristic approach used the perception of experts in the field to quantify the relationships between physical deterioration of assets and their properties. The perception of experts is collected through a survey, in which the evaluation of the influence of each factor on pipe deterioration is evaluated.

The most relevant factors that affect the pipe deterioration are organised in external and internal factors (presented in Table 6-1) and used to develop a survey for the field experts. The survey was developed online through the platform Google Forms as a means to efficiently collect and process the obtained answers from the experts and is presented in Appendix A. It was recommended that experts avoided guesswork as inaccurate results could be obtained.

Table 6-1 – Factors surveyed.

Type of factor	Factor
Internal	Material, age, nominal diameter, length, wall thickness, roughness coefficient, laying condition, average operating temperature, pressure variation,
External	Type of soil, soil density, backfill material, groundwater level, road traffic, type of road surface (asphalt, seal, unpaved), Type of area (residential, commercial, industrial), soil pH, soil resistivity, frost load, manufacturing process for metallic pipes (cast/spun), Cathodic protection for metallic pipes.

A total of 31 experts were surveyed and their experience distribution is presented in Figure 6-1. The number of experts in each experience group is similar, allowing to be obtained fairly distributed results throughout experience groups.

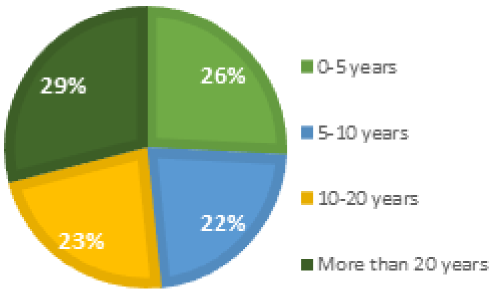


Figure 6-1 – Surveyed experts experience distribution: 0-5 years; 5-10 years; 10-20 years; more than 20 years.

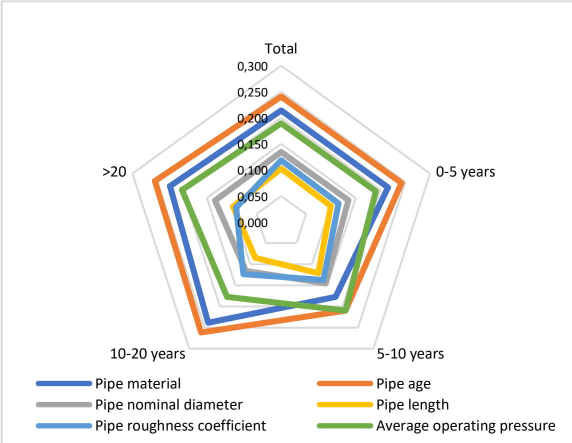
6.3 Model construction

The second step in the heuristic approach is the processing and analysis of the survey results to obtain quantitative weights for each influencing factor. The mathematical formulas developed to obtain the weights are described in Table 3-3. Each formula aims to value a certain aspect of the collected data that allows the comparison of the influence of each factor. For example, the use of average versus median values allows to evaluate the susceptibility to outliers in the data collected. The obtained weights are then normalized and presented in Appendix E.

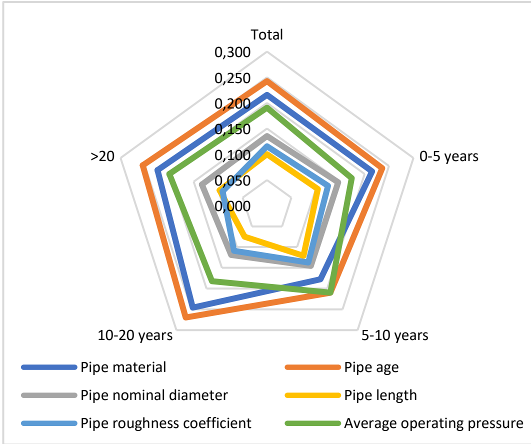
However, as the available data in the case study does not exist for all considered factors, only part of these were analysed. The factors which are available are: pipe material, pipe age, pipe nominal diameter, pipe length, pipe roughness coefficient and average operating pressure, which is considered the same as the static pressure for the purpose of this study. The respective normalized weights were recalculated accordingly. The weights associated with each factor calculated by each mathematical formulation (see Table 3-3) is presented in Figure 6-2. Note that the vertices of the diagram represent the weight attributed by each expert experience group.

The pipe age is the most relevant factor on pipe deterioration. However, the experience group of 5-10 years also considers the average operating pressure of pipes as an equally important influencing factor of deterioration for the majority of the weight formulas. On the other hand, the evaluation of the weight associated with each factor is very similar throughout experience groups with the order of most influential to least influential factors: age, material, average operating pressure, nominal diameter, roughness coefficient and length. Although some variation of this order can be identified, mainly in Figure 6-2(e), the overall consensus of the experts remains the former.

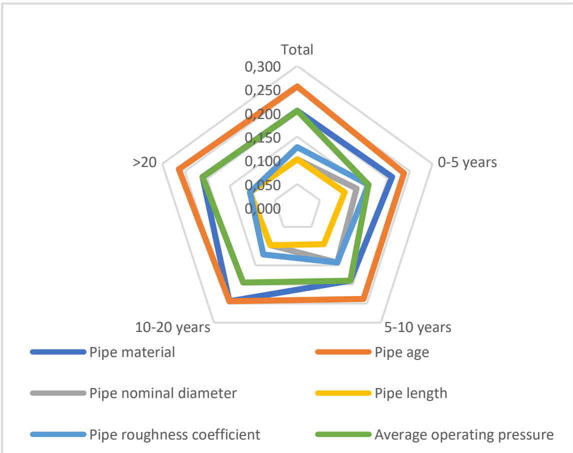
Besides, it can also be concluded that the order of influence in the factors does not change with the different weigh calculation methods. Furthermore, the multiplication of the number of replies did not significantly change the results, which can be explained by the recommendation for experts to avoid guess work in the survey leading to potential outliers in the answers.



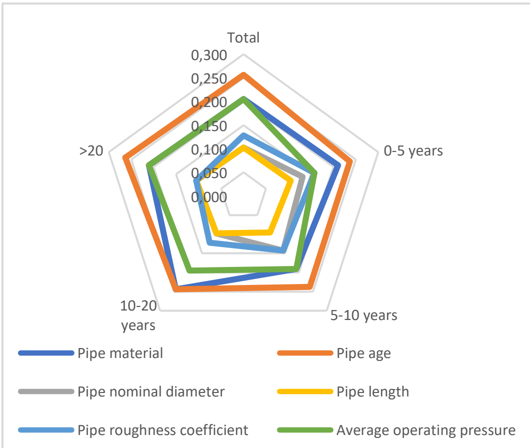
(a)



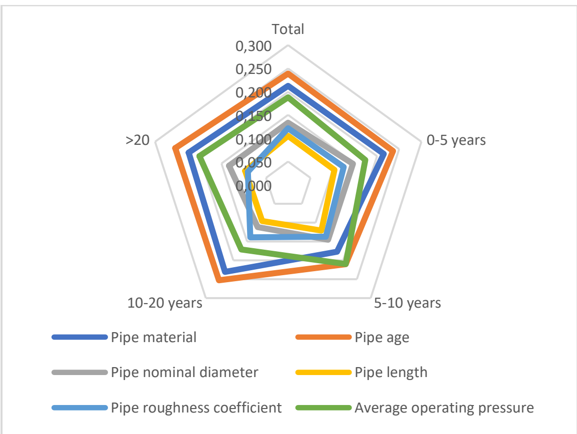
(b)



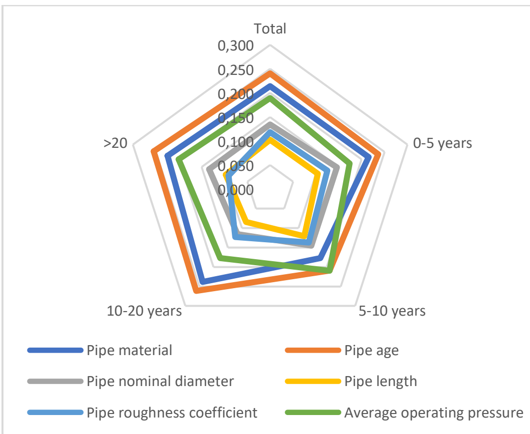
(c)



(d)



(e)



(f)

Figure 6-2 – Weight calculation for each influencing factor considering different weight formulas: (a) Mean value; (b) Mean * # Replies; (c) Median; (d) Median * # Replies; (e) Geometric average; (f) Geometric average * # Replies.

Nevertheless, when analysing the weights attributed by expert experience groups, different perceptions arise. While all expert experience groups agree that pipe age is highly influential, the experience group of 10 – 20 years concludes that pipe material and pipe age are highly influential, while the experts with more than 20 years of experience attribute a higher influence to average operating pressure than to pipe material.

Table 6-2 presents the results of the weight values using simple weight and the total expert knowledge, in which the evaluation of a pipe condition is carried out by:

$$C = 0.2407X_1 + 0.21415X_2 + 0.13451X_3 + 0.11858X_4 + 0.10265X_5 + 0.18938X_6 \qquad (6.1)$$

where the X_i represents the value associated with the pipe property *i*.

Table 6-2 – Weights associated with simple weight sum, total expert knowledge.

Factor	**Pipe age (X_1)**	**Pipe material (X_2)**	**Pipe nominal diameter (X_3)**	**Pipe roughness coefficient (X_4)**	**Pipe length (X_5)**	**Average operating pressure (X_6)**
Weight	0.2407	0.21415	0.13451	0.11858	0.10265	0.18938

6.4 Development of condition rating scale

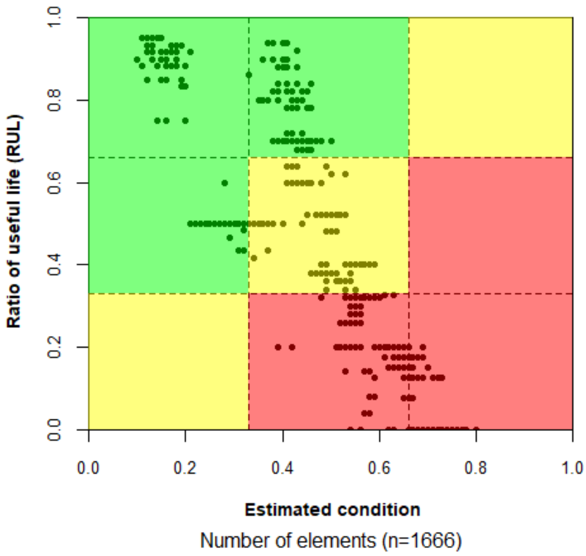
In both the deterministic and heuristic approach, the RUL is the variable chosen for the y-y axis of the condition assessment matrix. In the heuristic approach the variable closed for the xx- axis is the obtained pipe condition variable (i.e. the weighted sum) which is divided in three equal parts between the minimum and the maximum value. The condition scale developed is presented in Figure 3-2(a).

6.5 Condition assessment

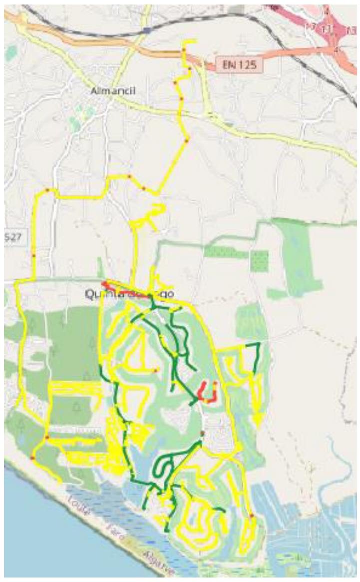
The application of the final step of the condition assessment using a heuristic approach to model construction requires the normalization of the properties of each pipe in the network according to Equation 3.8.

Figure 6-3 and Figure 6-4 present the results of two models obtained by the heuristic approach: considering the 31 experts and only considering the most experience experts (over 20-year experience). Obtained results are quite similar, classifying most of the network with an average value of condition.

Table 6-3 presents the comparison between the two condition assessments presented. The remaining analysis is presented in Appendix F. Table 6-3 demonstrates that the heuristic approach evaluates the condition of the current water distribution network mostly in an average condition.

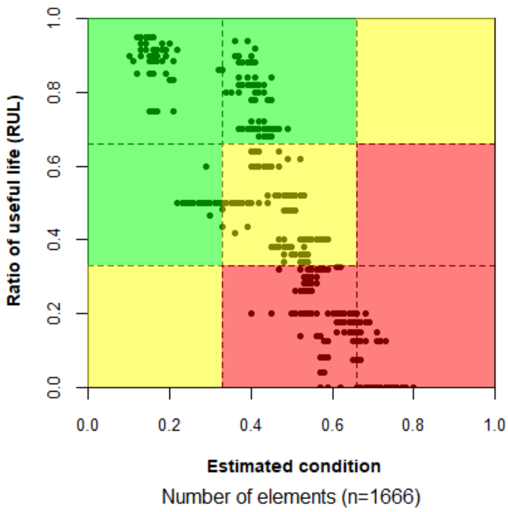


(a)

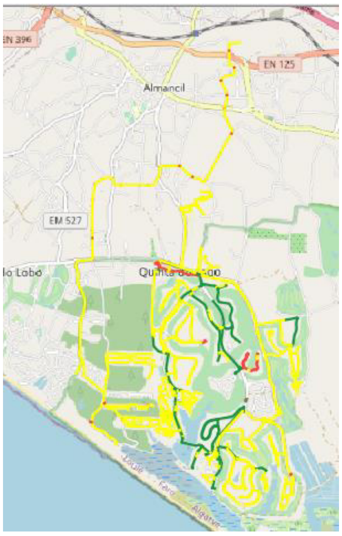


(b).

Figure 6-3 – Heuristic approach results using simple weight and all expert: (a) Network condition assessment with condition assessment matrix; (b) Map of network with designated condition.



(a)



(b)

Figure 6-4 – Heuristic approach results using simple weight and experts with more than 20 years of expertise (a) Network condition assessment with condition assessment matrix; (b) Map of network with designated condition.

Table 6-3 – Heuristic results for model summary: simple weight sum for total expert group and +20 expert group.

Condition grade	**Simple weight sum, total expert group**		**Simple weight sum, +20 expert group**	
	Number of pipes	**(%)**	**Number of pipes**	**(%)**
Good	370	22.21%	201	12.06%
Average	994	59.66%	1165	69.93%
Unsatisfactory	302	18.13%	300	18.01%
Total	**1666**	**100%**	**1666**	**100%**

6.6 Summary and conclusions

The heuristic approach is a knowledge-based approach that aims to take advantage of the existing state-of-the-art and to establish relationships between network properties and the physical condition of water distribution pipes. The main advantage of this approach is that survey results can be applied in other case studies. Consequently, the weights attributed by experts in the model construction are directly applicable to other distribution networks.

The broad analysis of the results of the heuristic approach allows to conclude that results do not present significant variation regarding the experience group and the weight formulas. These conclusions are useful for future applications of this methodology as these demonstrate that it is not necessary to evaluate weights using different formulas and experts with high experience.

7. Final discussion

Firstly, the analysis of the results obtained by the deterministic and heuristic approaches demonstrates that the deterministic approach tends to overestimate the pipe condition rating that is to classify pipes as more deteriorated than they effectively are. The overall condition of the analysed network, *Quinta do Lago*, is classified as good based on the performance indicator rate of pipe rehabilitation, AA03 (Alegre *et al.*, 2021), thus, it would be expectable that most of the pipes would be classified as good according to the proposed approaches. However, none of approaches (deterministic and heuristic) corroborate with this assessment and generally classified the pipes in worse condition.

Secondly, the application of the deterministic and heuristic approaches yielded different overall classifications for the same network pipes. While the heuristic approach classified 70% of the pipes with average condition and 18% unsatisfactory with bad condition (considering the linear model and the independent variables of average age at burst and static pressure), the deterministic approach evaluated approximately 60% of the pipes with unsatisfactory condition (considering simple weighted sum and the entire survey group). The difference in results can be attributed to the main assumption considered in the deterministic approach (see 5.6), which is that the physical condition of pipes without burst is similar to the condition of pipes with bursts (the original dataset). Consequently, a quite penalizing approach is attained with the deterministic models.

The development of a rehabilitation plan for a water distribution network should ideally integrate the results of both approaches. The results from both approaches are presented in Figure 5-6 and Figure 6-3. These should be compared to obtain the priority assets for intervention: the priority assets should be those classified in unsatisfactory condition in both approaches.

The cross-referencing of these two modelling approaches is described in Table 7-1. These results reduce the number of pipes classified in unsatisfactory condition to a total of 12. These 12 pipes should be considered the priorities for rehabilitation since an unsatisfactory condition is obtained in both approaches. These pipes do not have any relationship between each other in terms of location. However, when analysing the properties of these pipes, all pipes have the same material and the installation year between 1986 and 1990.

Table 7-1 – Cross-referencing of determinist and heuristic approaches.

Condition	Number of pipes with the same classification in the two approaches
Good	0
Average	178
Unsatisfactory	12
Total	**190**

Note: results of deterministic approach were obtained considering the linear model of average age at burst and static pressure and results of heuristic approach were obtained considering the simple weight average and the total survey group.

The assessment of the physical condition of water distribution pipes combining the well-known performance indicator (RUL) and new indicators (number of bursts per year for the deterministic approach and the non-dimensional condition variable for the heuristic approach) that are affected by the operation of the network to enhance the condition assessment tools available for water utilities. Previously the use of age as the primary factor to assess the pipes' condition yielded not very robust results.

The comparison of this previously condition assessment method (ratio of residual life) and the two new approaches is presented in Figure 7-1. Using a condition assessment matrix that combines results of models that consider different factors increases the robustness of the overall assessment methodology. The comparison of the results of the ratio of useful life with those of two approaches (Figure 7-1) demonstrates that the former tends to distribute pipe classification more equally in the three classes, the deterministic approach classifies most pipes in unsatisfactory condition and the heuristic approach in the average condition.

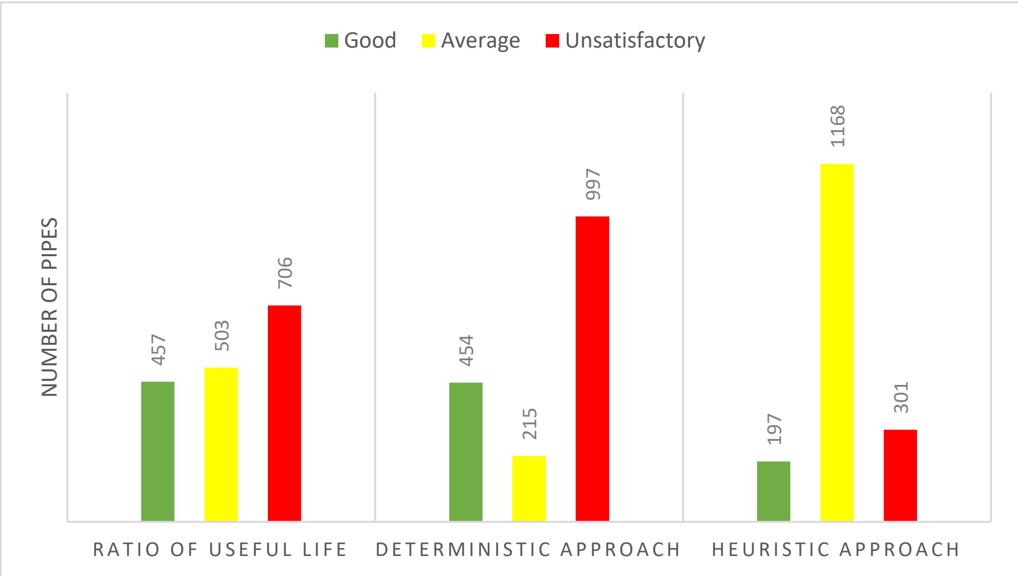
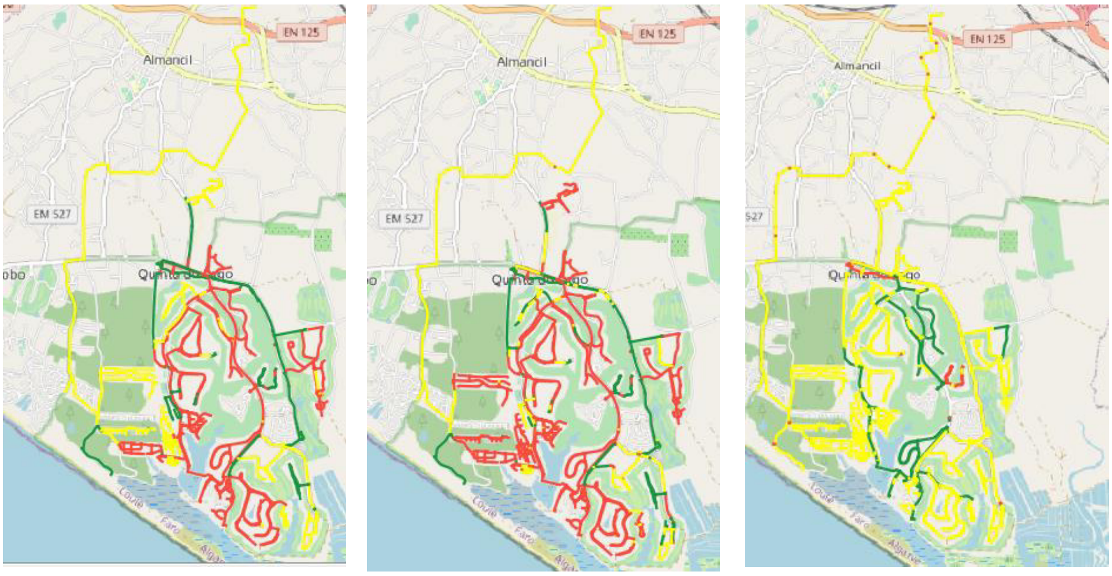


Figure 7-1 – Distribution of condition ratings for: (a) ratio of useful life; (b) deterministic approach; and (c) heuristic approach.

Figure 7-2 summarizes the condition assessment maps obtained by the referred three approaches: ratio of useful life, heuristic approach (simple weighted average) and deterministic approach (linear model of average age at burst and static pressure). The pipes classified in unsatisfactory conditions by the ratio of useful life and the deterministic approach are almost the same, though the heuristic approach is less conservative with a higher number of pipes in average condition.



(a) (b) (c)

Figure 7-2 – Colour grid map for: (a) ratio of useful life; (b) deterministic approach – linear model of average age at burst and static pressure; and (c) heuristic approach - simple weighted average.

The use of deterministic and heuristic approaches for pipe condition assessment of water distribution systems has significantly improved the existing methods, since these allow to include several factors that affect the pipe deterioration as well as the use of advanced statistical tools to model data. The proposed approaches should be incorporated in the infrastructure asset management of urban water assets contributing to more informed and robust decision-making processes.

8. Conclusions and future works

Worldwide, the number of water utilities that have operational profits is very small and those who have are often due to their specific financial conditions (Pinsent Masons LLP, 2011). Consequently, the supply of drinking water to domestic households is predominantly a non-self-sustaining economic activity mostly due to the insufficient rehabilitation rates that cannot be supported by the applied water tariffs. As the increase of water price is not considered a viable option, since the access to potable water is considered a human right, the optimization of the use of human, technological and financial resources can be the solution to decrease the financial problems of water utilities.

The improvement of condition assessment methods is essential to improve the rehabilitation efforts of assets and maximise the use of reduced budgets. This process can be carried out using direct and indirect condition assessment methods, however these are often too complex for water utilities to apply, since these require extensive fieldwork (i.e., direct condition assessment methods) or specific computational know-how (i.e., indirect condition assessment methods).

The proposed methodology allows the assessment of the physical condition of assets overcoming the described issues. The use of readily and available computing programs (i.e., Excel and RStudio) allows water utilities to easily assess the physical condition of water distribution networks without a high initial investment based on the knowledge of simple infrastructure and operational data (e.g., pipe characteristics and operating pressures).

A four-step methodology was proposed including data collection and validation, model construction using both deterministic and heuristic approaches, development of a condition rating scale appropriate for each approach and, finally, the condition assessment of the network assets. The application of the proposed methodology allowed to assess the physical condition of the *Quinta do Lago* water distribution network located in the district of the Algarve, Portugal.

Results from the two approaches presented significant differences in the overall pipe condition classification. The deterministic approach uses as a training dataset the pipes with burst during the period of observation. Consequently, the developed models include the notion that the structural condition of all other pipes will resemble the pipes that are presumably in worse condition (i.e., pipes that exhibited bursts during the period of observation). Therefore, this approach is considered to be highly penalizing in the classification of network pipes. On the other hand, the results presented by the heuristic approach were much less conservative, classifying the pipes as predominantly in average condition.

The proposed approaches can be considered an improvement on the existing methods of condition assessment through the use of the performance indicators, such as the indicator rehabilitation rate or the ratio of useful life. This improvement is justified by the incorporation of several factors that influence pipe deterioration and the association of preliminary network evaluation through the indicator RUL.

The future works should include the application of the proposed methodology to a larger dataset with a higher period of observation. This would allow to achieve more robust results, as well as to test the proposed methodology. Furthermore, results obtained should be included in the decision-making process on pipe rehabilitation.

On the other hand, the improvement of the proposed methodology can be achieved by employing more efficient, accurate and cost-effective approaches for predicting and explaining the pipe condition, such as, data mining techniques (e.g., artificial neural networks, decision trees). Although the development of straightforward approaches is one of the goals of this study, the use of stochastic approaches through an easy-to-use interface, such as a plugin in a geographic information system can increase accessibility and robustness of the current methodology.

9. References

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. E., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, *4*(11), e00938. https://doi.org/10.1016/j.heliyon.2018.e00938

Achim, D., Ghotb, F., & McManus, K. J. (2007). Prediction of water pipe asset life using neural networks. *Journal of Infrastructure Systems*, *13*(1), 26–30. https://doi.org/10.1061/(ASCE)1076-0342(2007)13:1(26)

Adorno, L., Hampshire, N., Cityworks, G. B., & Epcor, A. C. (2018). *AWWA Asset Management Definitions Guidebook* (1st ed., Issue January, pp. 1–33). American Water works association. https://www.awwa.org/Portals/0/AWWA/ETS/Resources/AMGuidebook.pdf?ver=2018-12-13-100101-887

Agarwal, M. (2010). Developing a Framework for Selecting Condition Assessment Technologies for Water and Wastewater Pipes. In *Virginia Polytechnic Institute and State University* (Issue 1).

Al-Barqawi, H., & Zayed, T. (2006a). Assessment model of water main conditions. *Proceedings of the 2006 Pipeline Division Specialty Conference - Pipelines 2006: Service to the Owner*, *211 40854*, 27. https://doi.org/10.1061/40854(211)27

Al-Barqawi, H., & Zayed, T. (2006b). Condition rating model for underground infrastructure sustainable water mains. *Journal of Performance of Constructed Facilities*, *20*(2), 126–135. https://doi.org/10.1061/(ASCE)0887-3828(2006)20:2(126)

Al-Barqawi, H., & Zayed, T. (2008). Infrastructure management: Integrated AHP/ANN model to evaluate municipal water mains' performance. *Journal of Infrastructure Systems*, *14*(4), 305–318. https://doi.org/10.1061/(ASCE)1076-0342(2008)14:4(305)

Alegre, H. (2008). *Gestão patrimonial de infra-estruturas de abastecimento de água e de drenagem e tratamento de águas residuais* (LNEC (ed.); 1st ed.). LNEC. www.lnec.pt

Alegre, H., & Coelho, S. (2012). Infrastructure Asset Management of Urban Water Systems. *Water Supply System Analysis - Selected Topics*. https://doi.org/10.5772/52377

Alegre, H., & Covas, D. (2010). Gestão patrimonial de infra-estruturas de abastecimento de água Uma. In E. R. dos S. de Á. e R. L. N. de E. C. I. S. Técnico (Ed.), *Série Guias Técnicos* (Vol. 16, Issue 1). https://doi.org/10.1071/MU949071

Alegre, H., Matos, R., Beja Neves, E., Cardoso, A., & Duarte, P. (2021). Guia de Avaliação da Qualidade dos Serviços de Águas e Resíduos Prestados aos Utilizadores. In Instituto Regulador de Águas e Resíduos (Ed.), *Série Guias Técnicos* (3rd ed., Vol. 22, Issue 3). Europress, Lda. http://www.ersar.pt/pt/publicacoes/publicacoes-tecnicas/guias

Alegre, H., Vitorino, D., & Coelho, S. (2014). Infrastructure value index: A powerful modelling tool for combined long-term planning of linear and vertical assets. *Procedia Engineering*, *89*, 1428–1436. https://doi.org/10.1016/j.proeng.2014.11.469

Amani, N., Nasly, M. A., & Samat, R. A. (2012). Infrastructure Component Assessment Using the Condition Index System: Literature Review and Discussion. *Journal of Construction Engineering and Project Management*, *2*(1), 27–34. https://doi.org/10.6106/jcepm.2012.2.1.027

American Water Works Service Co., I. (2002). Deteriorating Buried Infrastructure Management Challenges and Strategies. *Environmental Protection Agency (EPA)*, 1–33. http://www.epa.gov/ogwdw/disinfection/tcr/pdfs/whitepaper_tcr_infrastructure.pdf

Andreou, S. A., Marks, D. H., & Clark, R. M. (1987). A new methodology for modelling break failure patterns in deteriorating water distribution systems: Theory. *Advances in Water Resources*, *10*(1), 2–10. https://doi.org/10.1016/0309-1708(87)90002-9

API -Advanced pipe inspection UK Ltd. (2021). *Live Water Pipe Inspection*. http://www.api-uk.com/api-services/live-water-pipe-inspection/

Arbabi, H. (2005). *Corrosive Soils*. Testing Engineers. http://www.testing-engineers.com/case1.html

Arrival 3D. (2020). *UNDERSTANDING THE COST OF 3D SCANNING SERVICES*. https://arrival3d.com/3d-scanning-cost-variations-explained/

AWWA. (2014). Pipeline Renewal Methods. *M28 Rehabilitation of Water Mains, Third Edition*, *1*, 1–10.

Azevedo, L. F., Canário-Almeida, F., Fonseca, J. A., Costa-Pereira, A., Winck, J. C., & Hespanhol, V. (2011). Como escrever um artigo científico - estruturação e redacção da secção de métodos. *Revista Portuguesa de Pneumologia*, *17*(5), 232–238.

https://doi.org/10.1016/j.rppneu.2011.06.014
Babovic, V., Drécourt, J. P., Keijzer, M., & Friss Hansen, P. (2002). A data mining approach to modelling of water supply assets. *Urban Water*, *4*(4), 401–414. https://doi.org/10.1016/S1462-0758(02)00034-1
Bank, A. D. (2013). *Water Utility Asset Management: A Guide for Development Practioners* (1st ed.). Asian Development Bank. www.adb.org
Barton, N. A., Farewell, T. S., Hallett, S. H., & Acland, T. F. (2019). Improving pipe failure predictions: Factors effecting pipe failure in drinking water networks. *Water Research*, *164*. https://doi.org/10.1016/j.watres.2019.114926
Berardi, L., Giustolisi, O., Kapelan, Z., & Savic, D. A. (2008). Development of pipe deterioration models for water distribution systems using EPR. *Journal of Hydroinformatics*, *10*(2), 113–126. https://doi.org/10.2166/hydro.2008.012
Bishara, A. J., & Hittner, J. B. (2012). Testing the significance of a correlation with nonnormal data: Comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological Methods*, *17*(3), 399–417. https://doi.org/10.1037/a0028087
Blum, C., & Roli, A. (2003). Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison. *ACM Computing Surveys*, *35*(3), 268–308. https://doi.org/10.1145/937503.937505
Bravais, A. (1846). Sur les probabilit6s des erreurs de situation d'un point. In *Mermoirespresentes par divers savants a l'Academie royale des Sciencesde l'Institut de Ft'ance: Vol. IX*.
British Plastics Federation. (2021). *Hisotry*. https://www.bpfpipesgroup.com/about/history/
Bruaset, S., & Sægrov, S. (2018). An analysis of the potential impact of climate change on the structural reliability of drinking water pipes in cold climate regions. *Water (Switzerland)*, *10*(4). https://doi.org/10.3390/w10040411
Cabral, M., Loureiro, D., & Covas, D. (2019). Using economic asset valuation to meet rehabilitation priority needs in the water sector. *Urban Water Journal*, *16*(3), 205–214. https://doi.org/10.1080/1573062X.2019.1648528
Cabrera Rochera, E., Estruch-Juan, E., Gómez, E., & R. del Teso, R. (2019). cabrerar et al - Defining complementary tools to the IVI. The Infrastructure degradation index (IDI) and the Infrastructure Histogram (HI).pdf. *Urban Water Journal*. https://doi.org/https://doi.org/10.1080/1573062X.2019.1669195
Carmondy, S. (2017). Flint's mayor defends pipe replacement program. *Michigan Radio*. https://www.michiganradio.org/post/flints-mayor-defends-pipe-replacement-program
Cassa, A. M., van Zyl, J. E., & Laubscher, R. F. (2010). A numerical investigation into the effect of pressure on holes and cracks in water supply pipes. *Urban Water Journal*, *7*(2), 109–120. https://doi.org/10.1080/15730620903447613
Cassiolato, G., Carvalho, E. P., Caballero, J. A., & Ravagnani, M. A. S. S. (2021). Optimization of water distribution networks using a deterministic approach. *Engineering Optimization*, *53*(1), 107–124. https://doi.org/10.1080/0305215X.2019.1702980
Cawley, G. C., & Talbot, N. L. C. (2010). *On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation*. *11*, 2079–2107.
Clark, R. M., Carson, J., Thurnau, R. C., Krishnan, R., & Panguluri, S. (2010). Condition assessment modeling for distribution systems using shared frailty analysis. *Journal / American Water Works Association*, *102*(7), 81–91. https://doi.org/10.1002/j.1551-8833.2010.tb10151.x
Clark, R. M., Sivaganesan, M., Selvakumar, A., & Sethi, V. (2002). Cost Models for Water Supply Distribution Systems. *Journal of Water Resources Planning and Management*, *128*(5), 312–321. https://doi.org/10.1061/(asce)0733-9496(2002)128:5(312)
Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society Open Science*, *1*(3). https://doi.org/10.1098/rsos.140216
Covas, D., Cabral, M., Pinheiro, A., Marchionni, V., Antunes, S., Lopes, N., Mamouros, L., & Brôco, N. (2018). *Custos de construção de infraestruturas associadas ao ciclo urbano da água* (1st ed., Vol. 22, pp. 1–431). ERSAR. http://www.ersar.pt/pt/publicacoes/publicacoes-tecnicas/guias
Cox, L. A. (2008). What's wrong with risk matrices? *Risk Analysis*, *28*(2), 497–512. https://doi.org/10.1111/j.1539-6924.2008.01030.x
Csanyi, E. (2015). *4 Things Soil Resistivity Depends On*. Eletrical Engineering Portal. https://electrical-engineering-portal.com/4-things-soil-resistivity-depends-on
Dale, S. (2015). Heuristics and biases: The science of decision-making. *Business Information Review*, *32*(2), 93–99. https://doi.org/10.1177/0266382115592536

Direção-Geral de Saúde. (2020). *Amianto*. Saúde Pública. https://www.dgs.pt/saude-publica1/amianto.aspx
Decreto Lei 194/2009: Serviços municipais de abastecimento público de água, saneamento e resíduos urbanos, (2009).
Ministério Da Economia E Da Inovação, 5580 (2005) (testimony of DR n º199).
Limitação da Comercialização e Utilização do Amianto e dos Produtos que o Contenham, Diário da República 183 (1987).
DR nº219, 34012 (2016).
Duchesne, S., Bouzida, N., & Villeneuve, J.-P. (2011). Performance Estimation of a Remote Field Eddy Current Method for the Inspection of Water Distribution Pipes. *Journal of Water Resources Planning and Management*, *137*(6), 521–530. https://doi.org/10.1061/(asce)wr.1943-5452.0000136
Ennaouri, I., & Fuamba, M. (2013). New integrated condition-assessment model for combined storm-sewer systems. *Journal of Water Resources Planning and Management*, *139*(1), 53–64. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000217
Emergency administirve order, Section 1431 of the Safe Drinking Water Act (2016) (testimony of EPA).
European Union. (2020). *Portugal - Overview*. About the Eu. https://europa.eu/european-union/about-eu/countries/member-countries/portugal_en
Fahmy, M., & Moselhi, O. (2009). Forecasting the remaining useful life of cast iron water mains. *Journal of Performance of Constructed Facilities*, *23*(4), 269–275. https://doi.org/10.1061/(ASCE)0887-3828(2009)23:4(269)
Fares, H., & Zayed, T. (2010). Hierarchical Fuzzy Expert System for Risk of Failure of Water Mains. *Journal of Pipeline Systems Engineering and Practice*, *1*(1), 53–62. https://doi.org/10.1061/(asce)ps.1949-1204.0000037
Feeney, C., Thayer, S., Bonomo, M., & P.E, K. M. (2009). Condition Assessment of Wastewater Collection Systems. *Environmental Protection*, *May*.
Field, R. I., Murray, D. J., Royer, M. D., & Tafuri, A. N. (2007). Innovation and Research for Water Infrastructure for the 21St Century Research Plan. In *EPA* (Issue April). https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100ECPB.txt
Flores-Colen, I., Brito, J., & Peixoto de Freitas, V. P. (2008). Condition Assessment of Façade Rendering through in-Situ Testing. *11 DBMC International Conference on Durability of Building Materials and Components*. https://doi.org/10.13140/RG.2.1.3260.7441
Foulds, L. R. (2014). *The Heuristic Problem-Solving Approach Author (s): L . R . Foulds Source : The Journal of the Operational Research Society , Vol . 34 , No . 10 (Oct ., 1983), pp . 927- Published by : Palgrave Macmillan Journals on behalf of the Operational Research S*. *34*(10), 927–934.
Fukuyama, Y. (2008). K. Y. Lee and M.A. El-Sharkawi (editors), Modern Heuristic Optimization Techniques with Applications to Power Systems, IEEE Press Series on Power Engineering, John Wiley & Sons, 2008. In *Modern heuristic search methods*. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:MODERN+HEURISTIC+TECHNIQUES#3
Garvey, M. (2012). 3D laser scanning technology benefits pipeline design. *Pipeline and Gas Journal*, *239*(10), 1–6.
Gaspar, P. L. (2009). *Vida útil das construções: Desenvolvimento de uma metodologia para a estimativa da durabilidade de elementos da construção. Aplicação a rebocos de edifícios correntes*. Universidade de Lisboa- Instituto Superior Técnico.
Gottwald, S. (2020). Many-valued logic. In *Stanford Encyclopedia of Philosophy* (Summer 202, Issue 12). Metaphysics Research Lab, Stanford University. https://doi.org/10.5860/choice.41sup-0181
Green, H. (2019). *The Science of Flint's Water Crisis*. Scishow. https://www.youtube.com/watch?v=BAIXmt58iBU&ab_channel=SciShow
Harvey, R., de Lange, M., McBean, E., Trenouth, W., Singh, A., & James, P. (2017). Asset condition assessment of municipal drinking water, wastewater and stormwater systems–Challenges and directions forward. *Canadian Water Resources Journal*, *42*(2), 138–148. https://doi.org/10.1080/07011784.2016.1224684
Hassan Omer, N. (2020). Water Quality Parameters. *Water Quality - Science, Assessments and Policy*, 1–18. https://doi.org/10.5772/intechopen.89657
Howard, R. J. (2001). Infrastructure asset management under Australian Accounting Standard 27 (AAS27). *Proceedings of the Institution of Civil Engineers: Municipal Engineer*, *145*(4), 305–

310. https://doi.org/10.1680/muen.2001.145.4.305
Hu, Y., & Hubble, D. W. (2007). Factors contributing to the failure of asbestos cement water mains. *Canadian Journal of Civil Engineering*, *34*(5), 608–621. https://doi.org/10.1139/L06-162
Infraquinta. (2019). *Entidade Gestora : Ficha de avaliação da qualidade do serviço : Avaliação Valor do indicador Fiabilidade Histórico A entidade gestora deve promover um esforço de melhoria particularmente do (s) indicador (es) com avaliação insatisfatória . A entidade*.
IPWEA. (2006). Asset Performance Guidelines Practice Notes. *International Infrastructure Management Manual*, 1–26.
IPWEA. (2016). *Condtion assessment and asset performance guidelines*. *4*(1), 64–75.
ISO Guide 82:2019. (2019). *ISO 82:2019*.
J. Rumsey, D. (2016). *How to Determine a p -Value When Testing a Null Hypothesis*. Statistics Workbook For Dummies, Statistics II For Dummies, and Probability For Dummies; Wiley publishing Inc. https://www.dummies.com/education/math/statistics/how-to-determine-a-p-value-when-testing-a-null-hypothesis/
Jacobs, P., & Karney, B. (1994). GIS development with application to cast iron water main breakage rate. In B. G. Ltd (Ed.), *2nd international conference on water pipeline systems.*
Janela, J. M. E. M. (2017). *O Amianto Em Portugal. O Cumprimento Da Lei 2/2011, Sobre Amianto Em Edifícios Públicos*.
Jin Jun, H., Park, J. K., & Ho Bae, C. (2020). Economic Valuation of Aging Water Main Improvements. *Journal of Pipeline Systems Engineering and Practice*, *11*(1), 1–13. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000428
Kabir, G., Tesfamariam, S., Francisque, A., & Sadiq, R. (2015). Evaluating risk of water mains failure using a Bayesian belief network model. *European Journal of Operational Research*, *240*(1), 220–234. https://doi.org/10.1016/j.ejor.2014.06.033
Kayakan, E., & Khanesar, M. A. (2016). Fuzzy Neural Networks for Real Time Control Applications. In *Fuzzy Neural Networks for Real Time Control Applications*. Joe Hayton.
Kimutai, E., Betrie, G., Brander, R., Sadiq, R., & Tesfamariam, S. (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*, *6*(4), 04015005. https://doi.org/10.1061/(asce)ps.1949-1204.0000196
King, A. P., & Eckersley, R. J. (2019). Descriptive Statistics II: Bivariate and Multivariate Statistics. In *Statistics for Biomedical Engineers and Scientists* (pp. 23–56). https://doi.org/10.1016/b978-0-08-102939-8.00011-6
Krogh, A. (2008). What are artificial neural networks? *Nature Biotechnology*, *26*(2), 195–197. https://doi.org/10.1038/nbt1386
Laerd Statistics. (2020). *Pearson Product-Moment Correlation*. Statistical Tutorials and Software Guides. https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide-2.php
Lampola, T., & Kuikka, S. (2018). *Condition Assessment and Sewer Inspection (CASI) Methods – Guide Book* (Issue October).
Lawson, D., & Marion, G. (2008). *Introduction to Mathematical Modeling*. A First Course in Systems Biology. https://doi.org/10.4324/9780203702260-2
Lean, G. (1996). It took 80 years to act. *Independent*. https://www.independent.co.uk/news/uk/it-took-80-years-to-act-1303641.html
Liang, L. R., Lu, S., Wang, X., Lu, Y., Mandal, V., Patacsil, D., & Kumar, D. (2006). FM-test: A fuzzy-set-theory-based approach to differential gene expression data analysis. *BMC Bioinformatics*, *7*(SUPPL.4), 1–13. https://doi.org/10.1186/1471-2105-7-S4-S7
Liu, Z., & Kleiner, Y. (2012). State of the art review of inspection technologies for condition assessment of water pipes. *Measurement: Journal of the International Measurement Confederation*, *46*(1), 1–15. https://doi.org/10.1016/j.measurement.2012.05.032
Liu, Z., & Kleiner, Y. (2014). Computational intelligence for urban infrastructure condition assessment: Water transmission and distribution systems. *IEEE Sensors Journal*, *14*(12), 4122–4133. https://doi.org/10.1109/JSEN.2014.2336240
Lowe, M., & Cawley, P. (2006). *Long range guided wave inspection usage–current commercial capabilities and research directions*. 14–17. http://www3.imperial.ac.uk/pls/portallive/docs/1/55745699.PDF
Mahanta, J. (2017). *Introduction to Neural Networks, Advantages and Applications Artificial*. Towards Data Science. https://towardsdatascience.com/introduction-to-neural-networks-advantages-and-applications-96851bd1a207

Maji, A., & Jha, M. K. (2007). Modeling highway infrastructure maintenance schedules with budget constraints. *Transportation Research Record*, *1991*, 19–26. https://doi.org/10.3141/1991-03
Majumder, M. (2015). *Artificial Neural Network*. 49–54. https://doi.org/10.1007/978-981-4560-73-3_3
Makar, J. M., Desnoyers, R., & McDonald, S. E. (2001). Failure modes and mechanisms in gray cast iron pipes. *Underground Infrastructure Research*, 303–312. https://doi.org/10.1201/9781003077480-47
Marlow, D. R., & Burn, S. (2008). Effective use of condition assessment within asset management. *Journal / American Water Works Association*, *100*(1), 54–63. https://doi.org/10.1002/j.1551-8833.2008.tb08129.x
Martins, A. (2011). Stochastic models for prediction of pipe failures in water supply systems. *Stochastic Models for Prediction of Pipe Failures in Water Supply Systems*, 85. https://fenix.tecnico.ulisboa.pt/downloadFile/395143446036/Tese_Andre_Martins.pdf
Masten, S. J., Davies, S. H., & McElmurry, S. P. (2016). Flint water crisis: What happened and why? *Journal - American Water Works Association*, *108*(12), 22–34. https://doi.org/10.5942/jawwa.2016.108.0195
Matos, A. F. F. da S. C. (2015). *Modelação estatística da previsão da vida útil dos revestimentos de fachada* (Issue April) [Universidade de Lisboa- Instituto Superior Técnico]. https://doi.org/10.13140/RG.2.1.3667.0168
Mazumder, R. K., Salman, A. M., Li, Y., & Yu, X. (2018). Performance Evaluation of Water Distribution Systems and Asset Management. *Journal of Infrastructure Systems*, *24*(3), 03118001. https://doi.org/10.1061/(asce)is.1943-555x.0000426
Morgan, F. W., Kurtz, A. K., & Mayo, S. T. (1980). Statistical Methods in Education and Psychology. In *Journal of the American Statistical Association* (Vol. 75, Issue 371). https://doi.org/10.2307/2287695
Najafi, M., & Perez, M. (2016). Renewal of potable water pipes. In *Renewal of Potable Water Pipes* (Issue 132). https://doi.org/10.1061/9780784413746
National Research Council. (2003). *Deterioration and Inspection of Water Distribution Systems. 1.1*, 24.
Neter, J., J.Nachtsheim, C., & Kutner, M. H. (1983). Applied Linear Regression Models. In *Operations and Decision Sciences* (Fourth Edi, Vol. 26, Issue 4). McGraww-Hill Irwin. https://doi.org/10.2307/1269508
Novák, V., Perfilieva, I., & MocKor, J. (2013). Mathermatical principles of fuzzy logic. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9).
NSW. (2014). *NSW Reference Rates Manual. Valuation of water supply, sewerage and stormwater assets.*
Pearson, D., Fantozzi, M., Soares, D., & Waldron, T. (2005). Searching for N2: How does Pressure Reduction Reduce Burst Frequency? *Proceedings of the IWA Specialised Conference Leakage 2005*, 1–14.
Pure Tecnhologies. (2016). *SmartBall 2016 - Leak Detection - Animation*. https://www.youtube.com/watch?v=sfoBAW2FwG8&ab_channel=PureTechnologies
Quinta do Lago. (2020). *História da Quinta do Lago*. https://www.quintadolago.com/pt/resort/historia/
Rajani, B., & Kleiner, Y. (2001a). Comprehensive review of structural deterioration of water mains: Physically based models. *Urban Water*, *3*(3), 151–164. https://doi.org/10.1016/S1462-0758(01)00032-2
Rajani, B., & Kleiner, Y. (2001b). Comprehensive review of structural deterioration of water mains: statistical models. *Urban Water*, *3*(3), 151–164. https://doi.org/10.1016/S1462-0758(01)00032-2
Rajani, B., & Kleiner, Y. (2004). Non-destructive inspection techniques to determine structural distress indicators in water mains. *NRC Publications Archive (NPArC)*, 1–20. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Non-destructive+inspection+techniques+to+determine+structural+distress+indicators+in+water+mains#0
Rajani, B., & Makar, J. (2000). A methodology to estimate remaining service life of grey cast iron water mains. *Canadian Journal of Civil Engineering*, *27*(6), 1259–1272. https://doi.org/10.1139/l00-073
Rajeev, P., Kodikara, J., Robert, D., Zeman, P., & Rajani, B. (2014a). Factors contributing to large daimeter water pipe failure. *Water Asset Management International*, *10*(3), 9–14.
Rajeev, P., Kodikara, J., Robert, D., Zeman, P., & Rajani, B. (2014b). Factors contributing to large

daimeter water pipe failure. *Water Asset Management International*, *10*(3), 9–14.
Ramzai, J. (2020). *Clearly explained: Pearson V/S Spearman correlation coefficient*. Towards Data Science. https://doi.org/10.1016/j.ebobgyn.2004.04.015
RASARP. (2020). *Relatório anual dos Serviços de Águas e Resíduos em Portugal - 2019* (Vol. 1). Entidade Reguladora dos Servições de Águas e Resíduos.
RASARP. (2021). *Relatório anual dos Serviços de Águas e Resíduos em Portugal - 2020* (2020th ed., Vol. 1). Entidade Reguladora dos Servições de Águas e Resíduos.
Reddy, P. (2018). *The Differences Between Research Methods and Research Methodology*. The Diffrence Bewteen.Net. http://www.differencebetween.net/science/the-differences-between-research-methods-and-research-methodology/
Renaud, E., Bremond, B., & Le Gat, Y. (2014). Water pipes: Why 'lifetime' is not an adequate concept on which to base pipe renewal strategies. *Water Practice and Technology*, *9*(3), 307–315. https://doi.org/10.2166/wpt.2014.032
Rezaei, H., Ryan, B., & Stoianov, I. (2015). Pipe failure analysis and impact of dynamic hydraulic conditions in water supply networks. *Procedia Engineering*, *119*(1), 253–262. https://doi.org/10.1016/j.proeng.2015.08.883
Rizzo, P. (2010). Water and wastewater pipe nondestructive evaluation and health monitoring: A review. *Advances in Civil Engineering*, *2010*. https://doi.org/10.1155/2010/818597
Rose, J. L. (2003). Dispersion curves in guided wave testing. In *Materials Evaluation* (Vol. 61, Issue 1, pp. 20–22).
Royer, M. (2012). Condition assessment technologies for water transmission and distribution systems. *United States Environmental Protection Agency (EPA)*, *March*.
Ruchti, G. F. (2017). Water pipeline condition assessment. In *Water Pipeline Condition Assessment* (Issue 134). https://doi.org/10.1061/9780784414750
Selvakumar, A., Morisson, R., Sangster, T., Boyed Downey, D., & C.Matthews, J. (2013). *State of Technology for Rehabilitation of Water Distribution Systems* (Issue March).
Shi, Y., Zhang, C., Li, R., Cai, M., & Jia, G. (2015). Theory and application of magnetic flux leakage pipeline detection. *Sensors (Switzerland)*, *15*(12), 31036–31055. https://doi.org/10.3390/s151229845
Shoet, I. M., Rosenfeld, Y., Puterman, M., & Gilboa, E. (1999). Deterioration Patterns for Maintenance Management – a Methodological Approach. *Durability of Building Materials and Components*, *2 vols.*(8.), 1666–1678.
Singh, A., & Adachi, S. (2013). Bathtub curves and pipe prioritization based on failure rate. *Built Environment Project and Asset Management*, *3*(1), 105–122. https://doi.org/10.1108/BEPAM-11-2011-0027
St. Clair, A. M., & Sinha, S. (2012). State-of-the-technology review on water pipe condition, deterioration and failure rate prediction models! *Urban Water Journal*, *9*(2), 85–112. https://doi.org/10.1080/1573062X.2011.644566
Steven, B., & Jenkins, C. (1999). *Infrastructure Asset Grading Guidelines WATER ASSETS*.
Tai, C. C., Rose, J. H., & Moulder, J. C. (1996). Thickness and conductivity of metallic layers from pulsed eddy-current measurements. *Review of Scientific Instruments*, *67*(11), 3965–3972. https://doi.org/10.1063/1.1147300
Tavakoli, R. (2018). *Remaining Useful Life Prediction of Water Pipes Using Artificial* (Issue May). University of Texas.
Thomson, J., & Royer, M. (2013). *Primer on Condition Curves for Water Mains. December*.
Torrice, M. (2016). How lead eneded up in the water in flint michigan.pdf. *Chemical and Engineering News*. https://cen.acs.org/articles/94/i7/Lead-Ended-Flints-Tap-Water.html
Tuhovčák, L., Tauš, M., & Míka, P. (2014). Indirect condition assessment of water mains. *Procedia Engineering*, *70*(2008), 1669–1678. https://doi.org/10.1016/j.proeng.2014.02.184
Tuhovčák, L., Tauš, M., & Míka, P. (2013). Indirect condition assessment of water mains L. *12th International Conference on Computing and Control for the Water Industry*, 1669–1678.
U.S. Environmental Protection Agency. (2015). *Condition assessment of underground pipes. April*, 1–12. https://www3.epa.gov/region1/sso/pdfs/condition-assessment-underground-pipes.pdf
Ugarelli, R., & Bruaset, S. (2010). Review of deterioration modelling approaches for ageing infrastructure. *Journal of Chemical Information and Modeling*, *53*(9), 1689–1699.
Ugarelli, R., & Bruaset, S. (2013). Review of deterioration modelling approaches for ageing infrastructure. *SINTEF Rapport*, *53*(9), 1689–1699. https://doi.org/10.1017/CBO9781107415324.004
Ugarelli, R., Venkatesh, G., Brattebø, H., Di Federico, V., & Sægrov, S. (2010). Asset

Management for Urban Wastewater Pipeline Networks. *Journal of Infrastructure Systems*, *16*(2), 112–121. https://doi.org/10.1061/(asce)is.1943-555x.0000011
United Nations. (2015). TRANSFORMING OUR WORLD: THE 2030 AGENDA FOR SUSTAINABLE DEVELOPMENT. In *United Nations*. https://doi.org/10.1201/b20466-7
Urrea-Mallebrera, M., Altarejos-García, L., García-Bermejo, J., & Collado-López, B. (2019). Condition assessment of water infrastructures: Application to Segura River Basin (Spain). *Water (Switzerland)*, *11*(6), 1–21. https://doi.org/10.3390/w11061169
Uusitalo, L., Lehikoinen, A., Helle, I., & Myrberg, K. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling and Software*, *63*, 24–31. https://doi.org/10.1016/j.envsoft.2014.09.017
Vidigal, P. M., Covas, D., Loureiro, D., Coelho, S., & Alegre, H. (2010). Extensive analysis of hydraulic parameters in a large set of water distribution systems. *Integrating Water Systems - Proceedings of the 10th International on Computing and Control for the Water Industry, CCWI 2009*, 673–680.
Vieira, S., Pinaya, W. H. L., & Mechelli, A. (2017). Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neuroscience and Biobehavioral Reviews*, *74*, 58–75. https://doi.org/10.1016/j.neubiorev.2017.01.002
Wai-Lok Lai, W., Dérobert, X., & Annan, P. (2018). A review of Ground Penetrating Radar application in civil engineering: A 30-year journey from Locating and Testing to Imaging and Diagnosis. *NDT and E International*, *96*, 58–78. https://doi.org/10.1016/j.ndteint.2017.04.002
Wang, Y., Zayed, T., & Moselhi, O. (2009). Prediction models for annual break rates of water mains. *Journal of Performance of Constructed Facilities*, *23*(1), 40–46. https://doi.org/10.1061/(ASCE)0887-3828(2009)23:1(47)
Wood, A., & Lence, B. J. (2009). Using water main break data to improve asset management for small and medium utilities: District of Maple Ridge, B.C. *Journal of Infrastructure Systems*, *15*(2), 111–119. https://doi.org/10.1061/(ASCE)1076-0342(2009)15:2(111)
World Health Organization. (2000). Asbestos. In *Air quality guidllines* (2nd ed., Issue 2). https://www.euro.who.int/__data/assets/pdf_file/0015/123072/AQG2ndEd_6_2_asbestos.PDF
World Health Organization. (2014). Water Safety in Distribution Systems. *World Health Organization*, 157. http://apps.who.int/iris/bitstream/10665/204422/1/9789241548892_eng.pdf?ua=1
Woszczynski, M., Bergese, J., & Gagnon, G. A. (2013). Comparison of Chlorine and Chloramines on Lead Release from Copper Pipe Rigs. *Journal of Environmental Engineering*, *139*(8), 1099–1107. https://doi.org/10.1061/(asce)ee.1943-7870.0000712
Xylem. (2020). *SmartBall*. https://info.xyleminc.com/rs/198-DLL-407/images/SS_SmartBall_Platform_11052018_web.pdf
Zangenehmadar, Z. (2016). *Asset Management Tools for Sustainable Water Distribution Networks* (Issue August). Concordia University.
Zangenehmadar, Z., & Moselhi, O. (2015). *Application of Fahp and Shannon Entropy in Evaluating Criteria Significance in Pipeline Deterioration*. 1–10.
Zohuri, B., & Moghaddam, M. (2017). Business Resilience System (BRS): Driven through Boolean, Fuzzy logics and cloud computation: Real and near real time analysis and decision making system. *Business Resilience System (BRS): Driven Through Boolean, Fuzzy Logics and Cloud Computation: Real and Near Real Time Analysis and Decision Making System*, *March*, 1–425. https://doi.org/10.1007/978-3-319-53417-6

Appendices

Appendix A - Survey sent to experts in the urban water field

Influence of different factors in water distribution pipes' deterioration

Pipe characterisitic and operating conditions: Please select the importance of each factor listed below considering 0 not relevant and 5 very relevant. If you are not familiar with the factor select N/A *

	N/A	0	1	2	3	4	5
Pipe material	○	○	○	○	○	○	○
Pipe age	○	○	○	○	○	○	○
Pipe nominal diameter	○	○	○	○	○	○	○
Pipe length	○	○	○	○	○	○	○
Pipe wall thickness	○	○	○	○	○	○	○
Pipe roughness coefficient	○	○	○	○	○	○	○
Pipe laying conditions	○	○	○	○	○	○	○
Average operating pressure	○	○	○	○	○	○	○
Pressure variation	○	○	○	○	○	○	○

External factors #1: Please select the importance of each factor listed below considering 0 not relevant and 5 very relevant. If you are not familiar with the factor select N/A *

	N/A	0	1	2	3	4	5
Type of soil	○	○	○	○	○	○	○
Soil density	○	○	○	○	○	○	○
Backfill material	○	○	○	○	○	○	○
Backfill density	○	○	○	○	○	○	○
Groundwater level	○	○	○	○	○	○	○
Road traffic	○	○	○	○	○	○	○
Type of road surface (asphalt, seal or unpaved)	○	○	○	○	○	○	○
Type of area (residential, commercial, industrial)	○	○	○	○	○	○	○

External factors #2: Please select the importance of each factor listed below considering 0 not relevant and 5 very relevant. If you are not familiar with the factor select N/A *

	N/A	0	1	2	3	4	5
Soil pH	○	○	○	○	○	○	○
Frost load	○	○	○	○	○	○	○
Soil resistivity	○	○	○	○	○	○	○
Manufacturing process for metallic pipes (cast/spun)	○	○	○	○	○	○	○
Cathodic Protection for metallic pipes	○	○	○	○	○	○	○

Appendix B - Construction of linear models

Table B-1 – Linear models (n=49).

Y-Var	X-Var.	R^2	P-value	R^2_{adj}	Intercept	X_1	X_2	X_3	X_4	X_5	X_6
Nb urt/ YI	DT	0.05786	0.09593	0.03782	0.29130	-0.00002					
	Mat5	0.1055	0.02279	0.08646	0.58612	-0.00777					
	YI	0.1821	0.002227	0.1647	10.61837	-0.00523					
	SP	0.09571	0.03054	0.07647	0.45733	-0.00416					
	Length	0.05518	0.1042	0.03508	0.20510	0.00013					
	AA	0.3346	1.34E-05	0.3204	0.05821	0.00637					
	AA+DT	0.4443	1.35E-06	0.4201	0.12040	0.00693	-0.00003				
	AA+Mat5	0.3384	7.48E-05	0.3096	0.14377	0.00601	-0.00168				
	AA+YI	0.3637	3.06E-05	0.336	4.81394	0.00535	-0.00238				
	AA+SP	0.5062	8.93E-08	0.4848	0.33760	0.00717	-0.00565				
	AA+Length	0.3527	4.52E-05	0.3246	0.04786	0.00611	0.00008				
	DT+Mat3	0.1726	0.01281	0.1366	0.66180	-0.00002	-0.00812				
	DT+YI	0.2275	0.002641	0.1939	10.33000	-0.00002	-0.00506				
	DT+SP	0.097	0.09569	0.05774	0.48190	0.00001	-0.00491				
	DT+Length	0.09589	0.09841	0.05658	0.25730	-0.00002	0.00011				
	Mat3+YI	0.1877	0.008393	0.1523	9.27339	-0.00228	-0.00450				
	Mat3+SP	0.2122	0.004145	0.178	0.83892	-0.00818	-0.00440				
	Mat3+Length	0.125	0.04639	0.08693	0.51700	-0.00666	0.00008				
	YI+SP	0.2356	0.002074	0.2023	9.67770	-0.00467	-0.00317				
	YI+Length	0.2071	0.004815	0.1726	9.86600	-0.00486	0.00009				
	SP+Length	0.1109	0.06695	0.07225	0.40220	-0.00345	0.00008				
	AA+DT+Mat3	0.4475	5.96E-06	0.4106	0.19820	0.00660	-0.00003	-0.00153			
	AA+DT+YI	0.4586	3.82E-06	0.4225	3.49100	0.00617	-0.00003	-0.00169			
	AA+DT+SP	0.507	4.85E-07	0.4741	0.35670	0.00716	0.00000	-0.00623			
	AA+DT+Length	0.4487	5.68E-06	0.4119	0.11250	0.00677	-0.00003	0.00004			
	AA+Mat3+YI	0.3646	0.000126	0.3222	5.28266	0.00545	0.00097	-0.00264			
	AA+Mat3+SP	0.5086	4.52E-07	0.4758	0.40384	0.00688	-0.00132	-0.00563			
	AA+Mat3+Length	0.3534	0.0001839	0.3103	0.08690	0.00596	0.00075	0.00007			
	AA+YI+SP	0.5079	4.66E-07	0.4751	1.52630	0.00688	-0.00060	-0.00547			
	AA+YI+Length	0.377	8.22E-05	0.3354	4.42100	0.00520	-0.00219	0.00007			
	AA+SP+Length	0.5088	4.47E-07	0.4761	0.35870	0.00733	-0.00599	-0.00003			
	DT+Mat3+YI	0.2378	0.006271	0.187	8.45600	-0.00002	-0.00314	-0.00404			
	DT+Mat3+SP	0.213	0.01231	0.1605	0.85750	0.00000	-0.00816	-0.00499			
	DT+Mat3+Length	0.1803	0.02871	0.1257	0.61190	-0.00002	-0.00738	0.00005			
	DT+YI+SP	0.2366	0.006477	0.1858	9.80500	-0.00001	-0.00475	-0.00245			
	DT+YI+Length	0.2427	0.005469	0.1922	9.75900	-0.00002	-0.00478	0.00007			
	DT+SP+Length	0.1109	0.148	0.05164	0.40080	0.00000	-0.00341	0.00008			
	Mat3+YI+SP	0.2524	0.004161	0.2026	7.14223	-0.00409	-0.00329	-0.00358			
	Mat3+SP+Length	0.2122	0.01257	0.1597	0.84040	-0.00819	-0.00441	0.00000			
	Mat3+YI+Length	0.2083	0.01396	0.1555	9.26100	-0.00110	-0.00453	0.00008			
	YI+SP+Length	0.2416	0.00564	0.1911	9.40500	-0.00455	-0.00274	0.00005			
	AA+Mat3+YI+SP	0.5089	1.96E-06	0.4643	0.98969	0.00680	-0.00102	-0.00030	-0.00554		
	AA+Mat3+YI+Length	0.3806	0.0002482	0.3243	5.32000	0.00538	0.00197	-0.00269	0.00007		
	AA+DT+Mat3+YI	0.4586	1.52E-05	0.4094	3.59800	0.00619	-0.00003	0.00022	-0.00175		
	AA+Mat3+SP+Length	0.5132	1.63E-06	0.4689	0.45990	0.00698	-0.00186	-0.00609	-0.00004		
	AA+DT+Mat3+SP	0.5093	1.92E-06	0.4647	0.42160	0.00688	0.00000	-0.00130	-0.00619		
	Mat3+YI+SP+Length	0.2534	0.01062	0.1855	7.27200	-0.00369	-0.00337	-0.00336	0.00002		
	DT+Mat3+YI+SP	0.2527	0.01081	0.1848	7.24700	0.00000	-0.00402	-0.00335	-0.00321		
	DT+Mat3+YI+Length	0.2475	0.01237	0.1791	8.52500	-0.00002	-0.00223	-0.00411	0.00006		
	DT+Mat3+SP+Length	0.2131	0.02917	0.1416	0.87080	0.00001	-0.00827	-0.00514	-0.00001		
	AA+YI+SP+Length	0.5104	1.84E-06	0.4659	1.51200	0.00705	-0.00058	-0.00580	-0.00003		
	AA+DT+YI+SP	0.5083	2.01E-06	0.4636	1.41800	0.00691	0.00000	-0.00054	-0.00591		
	AA+DT+YI+Length	0.5083	2.01E-06	0.4636	1.41800	0.00691	0.00000	-0.00054	-0.00591		
	AA+DT+SP+Length	0.511	1.79E-06	0.4665	0.39810	0.00737	0.00001	-0.00710	-0.00004		
	DT+YI+SP+Length	0.2449	0.01321	0.1763	9.57700	-0.00001	-0.00466	-0.00134	0.00006		
	AA+DT+Mat3+YI+SP	0.5095	7.07E-06	0.4524	0.81560	0.00683	0.00000	-0.00110	-0.00020	-0.00607	
	AA+DT+Mat3+YI+Length	0.4622	4.47E-05	0.3997	3.69200	0.00612	-0.00003	0.00074	-0.00181	0.00004	
	AA+DT+Mat3+SP+Length	0.5158	5.44E-06	0.4595	0.50990	0.00701	0.00001	0.00196	-0.00733	-0.00006	
	AA+DT+YI+SP+Length	0.5119	6.41E-06	0.4551	1.28800	0.00714	0.00001	-0.00045	-0.00679	-0.00004	
	AA+Mat3+YI+SP+Length	0.5132	6.07E-06	0.4566	0.55920	0.00697	-0.00180	-0.00005	-0.00607	-0.00004	
	DT+Mat3+YI+SP+Length	0.2543	0.02301	0.1676	7.52700	-0.00001	-0.00341	-0.00352	-0.00256	0.00003	
	AA+DT+Mat3+YI+SP+Length	0.516	1.77E-05	0.4468	-0.01565	0.00709	0.00001	-0.00227	0.00028	-0.00756	-0.00006

Appendix C - Correlation matrices

Table C - 2 – Spearman's rank coefficient correlation matrix for data set n=49, only the data attributes related to clustering attempts.

	DCT	NB_Centroid	Mat_Cluster
Mat5	0.10	-0.26	0.01
Mat3	0.10	-0.26	0.01
YI	0.07	-0.28	0.07
YI_0	-0.07	0.28	-0.07
DN	-0.02	-0.10	0.15
Length	-0.19	0.16	0.24
AV_T	0.10	-0.13	-0.13
DT	-0.65	0.12	0.18
SP	-0.63	0.10	0.20
NB	-0.12	0.05	0.26
AA03	0.09	-0.09	-0.16
AA	-0.03	0.14	0.13
ROB	0.10	-0.09	-0.16
ROBLND	-0.12	0.05	0.26
NBY	-0.12	0.05	0.26
DCT	1.00	-0.56	-0.15
NB_Centroid	-0.56	1.00	-0.17
Mat_Cluster	-0.15	-0.17	1.00

- **DCT** – Distance from cluster centroid to tank
- **Nburst_Centroid** – number of bursts in the cluster.
- **Mat_cluster** – Dominant material type in the cluster.

Matrix legend

- **Mat5** – Material service life in five unique categories, see Table 5-2
- **Mat 3** – Material service life in three unique categories, see Table 5-3
- **YI** – Year of installation
- **YI_0** – Years in service
- **DN** – Nominal diameter
- **length** – Pipe length
- **AV_Temp** – Average water temperature at month of burst
- **DT** – Distance to tank
- **SP** – Average static pressure in the pipe
- **NB –** Number of bursts during observation period
- **AA03** – Rate of burst per 100 km [#/100km·year]
- **AA –** Average age of pipe at burst
- **ROB** – Rate of burst
- **ROBLND –** Rate of burst, where length is the average pipe length in the network
- **NBY–** Number of burst per year

Appendix D - Deterministic model results.

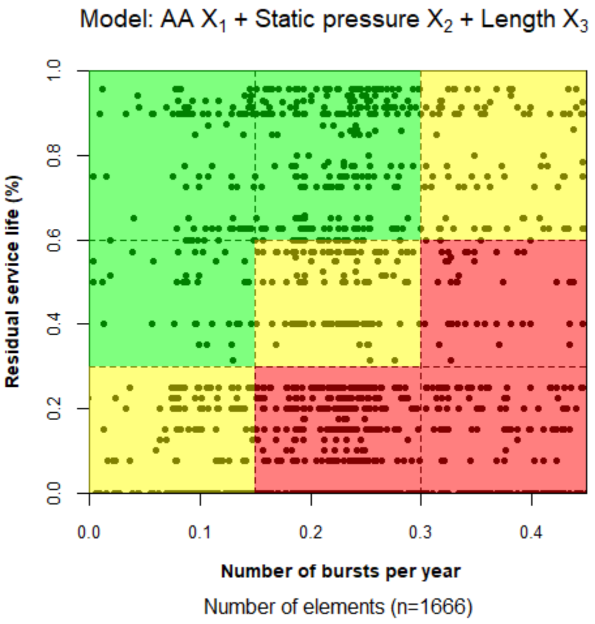


Figure D - 1 – Linear model AA+SP+Length.

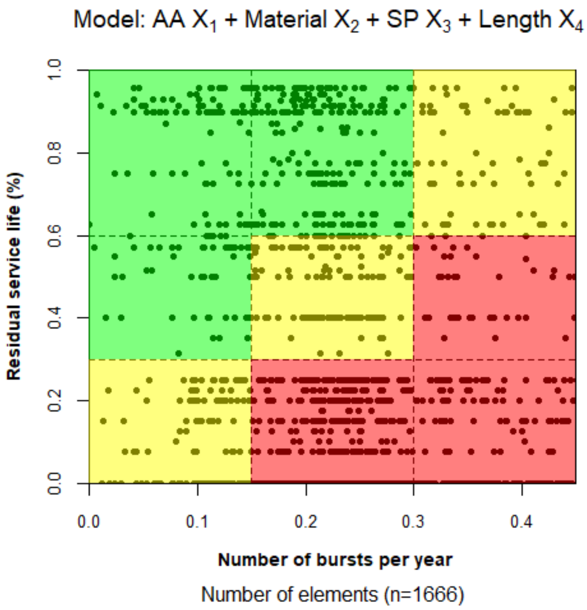


Figure D - 2 – Linear model AA + Material + SP +Length.

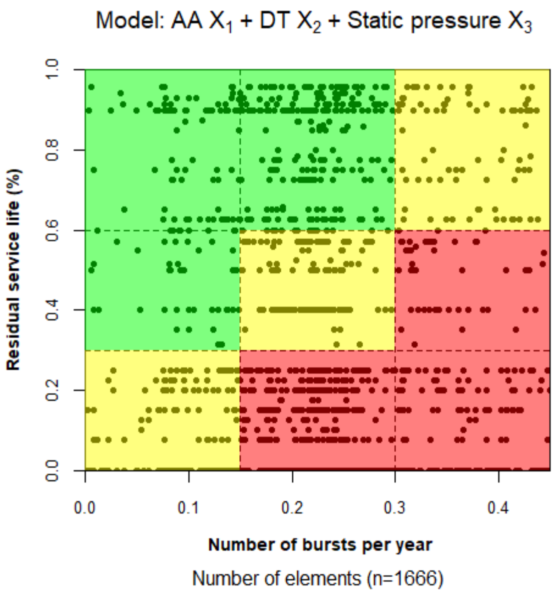


Figure D - 3 – Linear model AA+DT+SP.

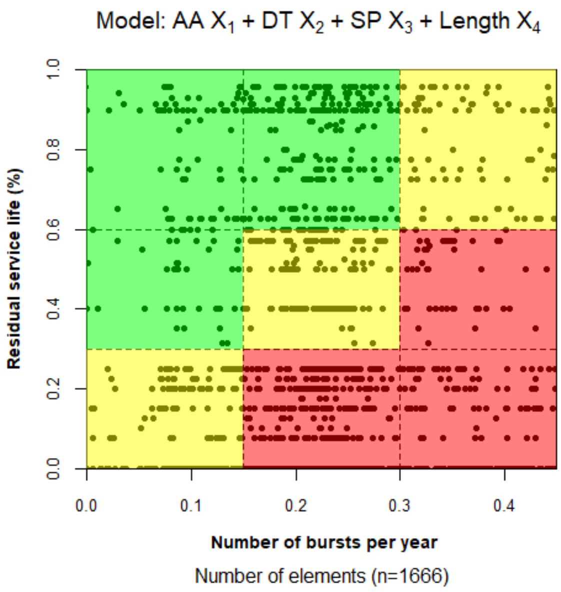


Figure D - 4 – Linear model AA + DT + SP + Length.

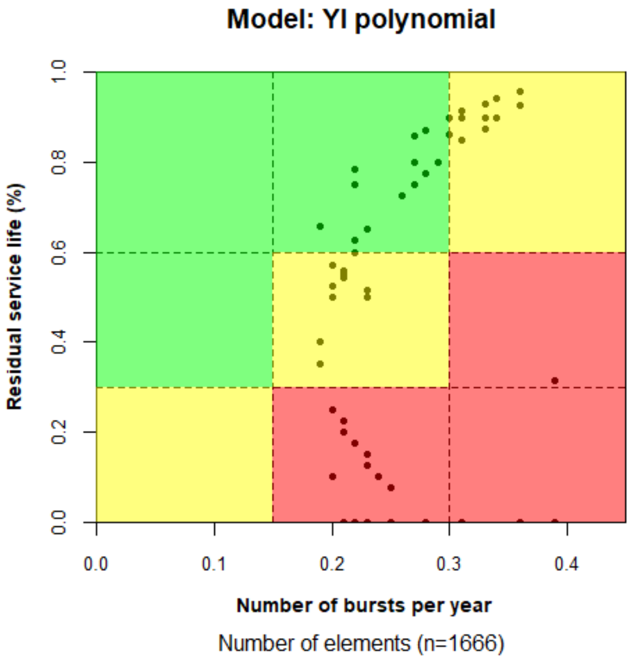


Figure D - 5 – Polynomial model for year of installation.

Appendix E – Survey results

Table E- 7 – Simple weight sum.

	Total	0-5 years	5-10 years	10-20 years	>20
Pipe material	0.061	0.065	0.061	0.0690	0.0608
Pipe age	0.069	0.076	0.072	0.0757	0.0690
Pipe nominal diameter	0.039	0.043	0.050	0.0334	0.0361
Pipe length	0.029	0.031	0.042	0.0245	0.0263
Pipe wall thickness	0.051	0.054	0.047	0.0535	0.0509
Pipe roughness coefficient	0.034	0.036	0.047	0.0356	0.0246
Pipe laying conditions	0.054	0.057	0.056	0.0535	0.0525
Average operating pressure	0.054	0.060	0.072	0.0512	0.0542
Pressure variation	0.067	0.073	0.075	0.0713	0.0608
Type of soil	0.052	0.052	0.042	0.0668	0.0509
Soil density	0.042	0.041	0.042	0.0490	0.0361
Backfill material	0.044	0.040	0.042	0.0490	0.0427
Backfill density	0.037	0.033	0.036	0.0290	0.0394
Groudwater level	0.046	0.046	0.033	0.0401	0.0427
Road Traffic	0.057	0.061	0.044	0.0557	0.0624
Type of road surface (asphalt, seal or unpaved)	0.045	0.048	0.039	0.0423	0.0411
Type of area (residential, commercial, industrial)	0.042	0.043	0.036	0.0379	0.0345
Soil pH	0.033	0.027	0.031	0.0223	0.0460
Frost load	0.035	0.027	0.036	0.0356	0.0312
Soil resistivity	0.027	0.020	0.036	0.0111	0.0361
Manufacturing process for metallic pipes (cast/spun)	0.039	0.030	0.036	0.0379	0.0427
Cathodic Protection for metallic pipes	0.045	0.033	0.025	0.0557	0.0591

Table E- 8 – Simple weight sum * Number of replies weight results.

	Total	0-5 years	5-10 years	10-20 years	>20
Pipe material	0.068	0.073	0.067	0.0746	0.0617
Pipe age	0.076	0.080	0.080	0.0818	0.0701
Pipe nominal diameter	0.043	0.050	0.055	0.0361	0.0367
Pipe length	0.031	0.035	0.046	0.0227	0.0267
Pipe wall thickness	0.054	0.056	0.052	0.0578	0.0517
Pipe roughness coefficient	0.036	0.043	0.052	0.0330	0.0250
Pipe laying conditions	0.057	0.060	0.061	0.0578	0.0534
Average operating pressure	0.060	0.059	0.080	0.0554	0.0551
Pressure variation	0.073	0.083	0.083	0.0770	0.0617
Type of soil	0.052	0.041	0.039	0.0722	0.0517
Soil density	0.041	0.035	0.039	0.0530	0.0367
Backfill material	0.040	0.027	0.039	0.0454	0.0434
Backfill density	0.033	0.024	0.034	0.0224	0.0400
Groudwater level	0.046	0.073	0.026	0.0371	0.0434
Road Traffic	0.061	0.078	0.042	0.0602	0.0634
Type of road surface (asphalt, seal or unpaved)	0.048	0.069	0.037	0.0457	0.0417
Type of area (residential, commercial, industrial)	0.043	0.069	0.034	0.0351	0.0350
Soil pH	0.027	0.009	0.029	0.0138	0.0467
Frost load	0.027	0.013	0.028	0.0330	0.0247
Soil resistivity	0.020	0.003	0.034	0.0052	0.0326
Manufacturing process for metallic pipes (cast/spun)	0.030	0.013	0.028	0.0292	0.0386
Cathodic Protection for metallic pipes	0.033	0.007	0.012	0.0516	0.0601

Table E- 9 – median value weight results.

	Total	0-5 years	5-10 years	10-20 years	>20
Pipe material	0.053	0.050	0.066	0.064	0.057
Pipe age	0.067	0.056	0.083	0.064	0.071
Pipe nominal diameter	0.027	0.031	0.050	0.026	0.028
Pipe length	0.027	0.025	0.033	0.026	0.028
Pipe wall thickness	0.053	0.050	0.050	0.051	0.057
Pipe roughness coefficient	0.033	0.038	0.050	0.032	0.028
Pipe laying conditions	0.053	0.050	0.050	0.051	0.043
Average operating pressure	0.053	0.038	0.066	0.051	0.057
Pressure variation	0.053	0.050	0.066	0.064	0.057
Type of soil	0.053	0.050	0.033	0.051	0.057
Soil density	0.040	0.044	0.033	0.038	0.028
Backfill material	0.040	0.050	0.041	0.045	0.043
Backfill density	0.040	0.050	0.025	0.038	0.028
Groudwater level	0.047	0.050	0.033	0.045	0.043
Road Traffic	0.053	0.050	0.041	0.051	0.057
Type of road surface (asphalt, seal or unpaved)	0.040	0.044	0.033	0.038	0.043
Type of area (residential, commercial, industrial)	0.040	0.050	0.033	0.045	0.043
Soil pH	0.040	0.044	0.025	0.038	0.043
Frost load	0.040	0.050	0.050	0.038	0.043
Soil resistivity	0.040	0.038	0.025	0.038	0.043
Manufacturing process for metallic pipes (cast/spun)	0.053	0.044	0.050	0.051	0.050
Cathodic Protection for metallic pipes	0.053	0.050	0.066	0.051	0.057

Table E- 10 – Median values * # replies weight results.

	Total	0-5 years	5-10 years	10-20 years	>20
Pipe material	0.060	0.062	0.075	0.072	0.057
Pipe age	0.075	0.070	0.094	0.072	0.071
Pipe nominal diameter	0.030	0.039	0.056	0.029	0.028
Pipe length	0.029	0.031	0.037	0.025	0.028
Pipe wall thickness	0.058	0.054	0.056	0.057	0.057
Pipe roughness coefficient	0.036	0.047	0.056	0.031	0.028
Pipe laying conditions	0.058	0.054	0.056	0.057	0.043
Average operating pressure	0.060	0.047	0.075	0.057	0.057
Pressure variation	0.060	0.062	0.075	0.072	0.057
Type of soil	0.054	0.047	0.032	0.057	0.057
Soil density	0.040	0.041	0.032	0.043	0.028
Backfill material	0.038	0.039	0.040	0.043	0.043
Backfill density	0.036	0.039	0.024	0.031	0.028
Groudwater level	0.047	0.062	0.027	0.043	0.043
Road Traffic	0.058	0.062	0.040	0.057	0.057
Type of road surface (asphalt, seal or unpaved)	0.043	0.054	0.032	0.043	0.043
Type of area (residential, commercial, industrial)	0.042	0.062	0.032	0.043	0.043
Soil pH	0.033	0.027	0.024	0.025	0.043
Frost load	0.032	0.031	0.040	0.037	0.043
Soil resistivity	0.029	0.018	0.024	0.018	0.043
Manufacturing process for metallic pipes (cast/spun)	0.042	0.027	0.040	0.041	0.050
Cathodic Protection for metallic pipes	0.040	0.023	0.032	0.049	0.057

Table E- 11 – Geometric average weight results.

	Total	0-5 years	5-10 years	10-20 years	>20
Pipe material	0.055	0.055	0.054	0.063	0.060
Pipe age	0.062	0.060	0.064	0.069	0.068
Pipe nominal diameter	0.034	0.037	0.044	0.030	0.035
Pipe length	0.027	0.026	0.037	0.026	0.026
Pipe wall thickness	0.047	0.055	0.042	0.048	0.050
Pipe roughness coefficient	0.031	0.032	0.042	0.038	0.024
Pipe laying conditions	0.050	0.059	0.049	0.048	0.052
Average operating pressure	0.048	0.044	0.064	0.046	0.053
Pressure variation	0.059	0.062	0.066	0.065	0.060
Type of soil	0.051	0.054	0.043	0.060	0.050
Soil density	0.041	0.047	0.043	0.044	0.035
Backfill material	0.046	0.051	0.043	0.052	0.042
Backfill density	0.040	0.045	0.037	0.037	0.039
Groudwater level	0.045	0.055	0.041	0.042	0.042
Road Traffic	0.053	0.058	0.046	0.050	0.061
Type of road surface (asphalt, seal or unpaved)	0.041	0.051	0.040	0.038	0.040
Type of area (residential, commercial, industrial)	0.040	0.051	0.037	0.040	0.034
Soil pH	0.040	0.028	0.031	0.035	0.045
Frost load	0.043	0.039	0.045	0.038	0.039
Soil resistivity	0.038	0.014	0.037	0.024	0.040
Manufacturing process for metallic pipes (cast/spun)	0.049	0.039	0.045	0.048	0.047
Cathodic Protection for metallic pipes	0.059	0.038	0.051	0.059	0.058

Table E- 12 – Geometric average * # replies weight results.

	Total	0-5 years	5-10 years	10-20 years	>20
Pipe material	0.061	0.065	0.061	0.069	0.061
Pipe age	0.069	0.072	0.072	0.076	0.069
Pipe nominal diameter	0.039	0.044	0.050	0.033	0.036
Pipe length	0.029	0.032	0.042	0.024	0.026
Pipe wall thickness	0.051	0.057	0.047	0.053	0.051
Pipe roughness coefficient	0.034	0.038	0.047	0.036	0.025
Pipe laying conditions	0.054	0.061	0.056	0.053	0.053
Average operating pressure	0.054	0.053	0.072	0.051	0.054
Pressure variation	0.067	0.074	0.075	0.071	0.061
Type of soil	0.052	0.048	0.042	0.067	0.051
Soil density	0.042	0.042	0.042	0.049	0.036
Backfill material	0.044	0.038	0.042	0.049	0.043
Backfill density	0.037	0.034	0.036	0.029	0.039
Groudwater level	0.046	0.065	0.033	0.040	0.043
Road Traffic	0.057	0.069	0.044	0.056	0.062
Type of road surface (asphalt, seal or unpaved)	0.045	0.061	0.039	0.042	0.041
Type of area (residential, commercial, industrial)	0.042	0.061	0.036	0.038	0.034
Soil pH	0.033	0.017	0.031	0.022	0.046
Frost load	0.035	0.023	0.036	0.036	0.031
Soil resistivity	0.027	0.006	0.036	0.011	0.036
Manufacturing process for metallic pipes (cast/spun)	0.039	0.023	0.036	0.038	0.043
Cathodic Protection for metallic pipes	0.045	0.017	0.025	0.056	0.059

Appendix F - Heuristic model results

Table F-1 – Results for heuristic model classification using total expert group.

Condition	Simple weight sum	Simple weight sum * #Replies	Median value	Median Value *#Replies	Normalised Average	Geometric Average *#Replies
Good	370	370	370	370	370	370
Average	994	994	994	994	994	994
Unsatisfactory	302	302	302	302	302	302

Table F-2 – Results for heuristic model classification using 0-5 expert group.

Condition	Simple weight sum	Simple weight sum * #Replies	Median value	Median Value *#Replies	Normalised Average	Geometric Average *#Replies
Good	**199**	**199**	**199**	**199**	**199**	**199**
Average	**1167**	**1167**	**1167**	**1167**	**1167**	**1167**
Unsatisfactory	**300**	**300**	**300**	**300**	**300**	**300**

Table F- 3 – Results for heuristic model classification using 5-10 expert group.

Condition	Simple weight sum	Simple weight sum * #Replies	Median value	Median Value *#Replies	Normalised Average	Geometric Average *#Replies
Good	199	199	199	199	199	199
Average	1167	1167	1167	1167	1167	1167
Unsatisfactory	300	300	300	300	300	300

Table F- 4 – Results for heuristic model classification using 10-20 expert group.

Condition	Simple weight sum	Simple weight sum * #Replies	Median value	Median Value *#Replies	Normalised Average	Geometric Average *#Replies
Good	370	370	370	370	370	370
Average	992	992	992	992	992	992
Unsatisfactory	304	304	304	304	304	304

Table F- 5 – Results for heuristic model classification using +20 expert group.

Condition	Simple weight sum	Simple weight sum * #Replies	Median value	Median Value *#Replies	Normalised Average	Geometric Average *#Replies
Good	201	201	201	201	201	201
Average	1165	1165	1165	1165	1165	2330
Unsatisfactory	300	300	300	300	300	900