# Assessing pipe condition in water distribution networks: deterministic versus heuristic models

Duarte Músrias Gray

duarte.gray@tecnico.ulisboa.pt

Department of Civil Engineering, Architecture and Georesources

Instituto Superior Técnico, Universidade de Lisboa

July 2021

## Abstract:

The condition assessment of water distribution pipes is of utmost importance for the prioritization of rehabilitation interventions. However, available methodologies for condition assessment are increasingly complex to be applied by water utilities with limited 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

## 1. Introduction

The concept *condition* can refer to the hydraulic capacity, the economic evaluation, the structural integrity or a combination of any of these factors (Thomson and Royer, 2013). In this study, this concept refers to the structural integrity of the evaluated pipes, also referred to as the pipe physical condition. Being 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, Murray, *et al.*, 2007).

The assessment of the physical condition of assets is an important process of infrastructure asset management (IAM), which corresponds to a set of processes that can be applied by water utilities to balance the three dimensions of analysis in the long-term – cost, risk and performance (Alegre and T.Coelho, 2012). An efficient IAM aims to provide an asset management plan (AMP), allowing to minimize the life-cycle costs by lowering renewal costs, identifying the necessary measures for optimized resource allocation and prioritizing the necessary interventions (Harvey, de Lange, *et al.*, 2017). Thus, a comprehensive methodology for IAM that considers decisions on strategic, tactical and operational levels for asset management is needed (Alegre and Covas, 2010). Condition assessment is integrated into IAM in the process of asset performance evaluation. The condition assessment of water distribution pipes can be carried out using direct or indirect condition assessment methods (Urrea-Mallebrera, Altarejos-García, *et al*., 2019). Direct condition assessment methods require the inspection of the asset, which 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).

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 to estimate the current physical condition of the asset (Liu and Kleiner, 2012; Al-Barqawi and Zayed, 2006b). These methods can be classified into 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 calculation) categories and aim to analyse trends and changes in the system to provide information on the condition of assets (Selvakumar, Morisson, *et al*., 2013; Ugarelli and Bruaset, 2013). 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 and Bruaset, 2013).

The development of indirect condition assessment methods is often based on the known relationships between deterioration and the factors that affect the physical condition of buried pipes, these factors can be divided into (Al-Barqawi and Zayed, 2006a):

- **Physical factors** – factors that are related to intrinsic pipe properties, such as pipe material, age, nominal 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 of the pipe, such as water quality, water pressure, backflow potential, flow velocity and operational and maintenance practices.

A burst in a water distribution pipe, if not caused by third-parties, can be attributed to pipe deterioration and interpreted as a sign of the decreased physical condition. This phenomenon can be attributed to a single influencing factor or a combination of factors. While the factors that contribute to the deterioration of water distribution pipes are known, the relative influence of each factor is widely discussed (Rajani and Kleiner, 2001). As a result, even though research has proven the existence of failure mechanisms, the uncertainty associated with the exact relationship between factors is still the focus of many studies. The main objective of this research is to develop a methodology to assess the physical condition of water distribution pipes, without the need for visual inspection or service interruption. The proposed methodology takes into consideration the knowledge gaps that Portuguese water utilities face, regarding the low infrastructure knowledge and the difficulty in using the current indirect condition assessment methods. Two approaches (deterministic and heuristic) are developed, compared and applied to assess the pipe physical condition of the water distribution network of Quinta do Lago.

## 2. Proposed methodology

The proposed methodology for the physical condition assessment of water distribution pipes is a four-step procedure (Figure 1). The application of this methodology aims to allow water utilities to efficiently assess the physical condition of their water distribution networks without the need for inspection or extensive field survey, which represents a less resource-consuming process. 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.

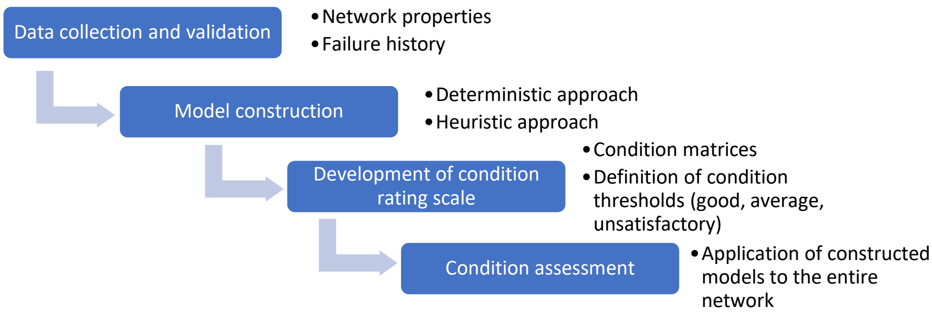



*Figure 1 – Proposed methodology for the physical condition assessment of water distribution pipes.*

The first step of the proposed methodology involves the collection of information regarding the factors that influence the physical condition of water pipes, namely network properties (e.g., pipe age, operating pressure, nominal diameter) and failure history (e.g., date of burst, location), also referred to as independent and dependent variables, respectively, in the model construction (Step 2). The collected factors that influence pipe deterioration can be divided into three main categories of physical, operational and environmental. However, the collection of the failure history is also required to assess the physical condition of assets without inspection. Table 1 presents the influencing factors of pipe deterioration which can be collected for the application of the proposed methodology.

*Table 1 – Potential influencing factors of 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. |

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

Although the factors described in Table 1 represent potential factors to assess the physical condition, some of these network properties are unknown for the majority of water utilities. Therefore, it is recommended that at least the variables presented in "bold" are collected as these are considered highly influential of pipe deterioration.

Before data analysis, the information gathered needs to be validated and checked for any inconsistencies (e.g., incorrect installation year of water pipes and inappropriate material choice for the nominal diameter). These inconsistencies can be found through a detailed study of the network, in which the analysis of network properties versus historical and geographic data reveals existing discrepancies. In most cases, these inconsistencies are the result of human error and can be corrected with the review of other network properties.

The second step of the proposed methodology is the construction of models, considering two mathematical modelling approaches: a deterministic modelling approach and a heuristic modelling 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 variable 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 dependent and independent variables used must be defined. The independent variables used are the properties of water distribution pipes (e.g., year of installation, material, nominal diameter) and the dependent variables are the failure history in the deterministic approach. Preferably, the collected variables should be quantitative, this operation might require the conversion of categorical variables to a numerical value.

An example of this conversion is the use of reference values of service lives to define the variable of pipe material. This association was based on the relevancy that material durability has with pipe deterioration. In this study, the established references values of service lives are based on Covas, Cabral, *et al*., (2018) and presented in Table 2.

*Table 2 – Established service lives.*

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

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 into the second step of the deterministic approach (i.e., regression model construction), as it allows for a preliminary understanding of the existing relationships in the case study network. The Spearman's rank correlation coefficient is used to construct the correlation matrix as it provides a nonparametric technique and measures the strength and direction of a non-monotonic relationship (King and Eckersley, 2019).

In this second 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 for both simple (r-squared) and multiple regression analysis (adjusted r-squared). In addition to the calculation of how well linear regressions adjusts 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, $\alpha$ , and rejecting the null hypothesis if the p-value is lower than $\alpha$ (Deborah J Rumsey, 2016). Finally, the last statistical analysis is to verify the independence between explanatory variables (i.e., independent variables). The mathematical measurement used to carried out this verification is the variance inflation factor (VIF). Values greater or equal to 10 are considered to have a high degree of multicollinearity and, therefore, are excluded from the models (Neter, J.Nachtsheim, *et al.*, 1983).

The third and final step of the deterministic approach is the validation of the linear models which involves verifying the statistical dispersion of the collected data, namely through the evaluation of the root mean square error, relative error or absolute error. The collected data must be divided into two subsets to carried out the validation: the training subset used for model construction, typically representing 80% of the whole dataset and the validation subset used to validate the obtained results through the training subset, representing the remaining 20% of the dataset. 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 and Talbot, 2010).

On the other hand, the heuristic approach relies on a theoretical framework for model construction, aiming to avoid the strict nature of some types of algorithms and achieve results. The main advantage of the heuristic approach is the ability to apply intuition to problem solving and reduce the resources needed to achieve a satisfactory result (Foulds, 2014).

The steps to complete the heuristic approach are threefold:

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 the 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.

In the developed 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. To facilitate the comprehension of the questionnaire the factors are grouped into: (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 applied to value different aspects of the collected data, namely the discrepancies in the average value (i.e., the usage of the average and median values) and the overall confidence of experts in their perceptions (i.e., the bias of the number of replies for each factor). Furthermore, it should be noted that weights should range from zero to one and the total weight sum should be one.

In 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. However, when developing this normalization, it should be noted that not all variables have the same relationship (positive or negative) with pipe physical condition. This means that special attention must be given to combine the variation of the influencing factor and their contribution to the condition grade. This is done by revisiting the knowledge gathered regarding the factors that influence the deterioration of pipes and translating their deterioration effects into the normalization of the variables.

The third step of the proposed methodology aims to develop a comprehensive condition rating scale to interpret the results of the previously developed models and to assess the physical condition of water distribution pipes. A condition matrix with three condition rating levels is developed: good, average and unsatisfactory. For each approach (i.e., deterministic and heuristic) a different condition matrix is developed with the modelled values being combined with the performance indicator infrastructure value index (IVI). The assumption behind these matrices arises from the ambition to combine IVI with data related to observations (e.g., pipe burst) made during the operation of the network. 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, considering a division into three equal intervals of the indicator. 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 adoption of the original values of the IVI scale can be used, however, the condition 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. Thus, the established condition intervals are different for each approach. In the heuristic approach, the same strategy of the RUL scale is proposed as the interval of variation of the evaluated variable is divided into three equal intervals. The proposed condition matrix for the heuristic approach is presented in Figure 2(a).

In the case of the deterministic approach, the development of the x-x axis of the condition 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. Since the number of bursts was chosen as a dependent variable in the deterministic approach, the observation period is a key element to define the condition levels that a water utility considers as acceptable. In this case, the issue that needs to be addressed is "*what is the acceptable number of bursts given the observation period?*".

In the case of an observation period of 6 years (as considered in the 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 or more for pipes in unsatisfactory condition. Figure 2(b) presents the proposed condition matrix for the deterministic approach for the case study of *Quinta do Lago.* In the case of a higher observation period, the acceptable number of bursts might increase.

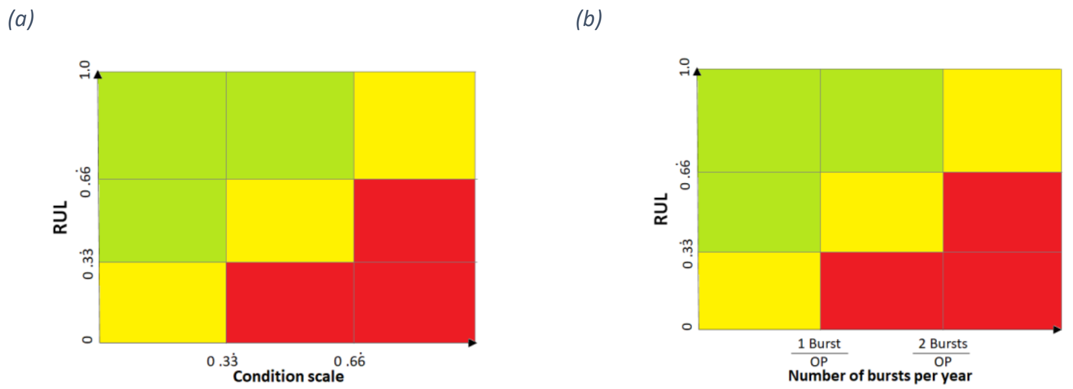



*Figure 2 - Condition matrices: (a) Heuristic approach; (b) Deterministic approach.*

## 3. Results

The water distribution network of *Quinta do Lago* is the case study of this research work. The data collected regarding network properties and 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 network. The distribution network has a total length of approximately 85 km and is constructed in different materials, such as polyvinyl chloride (PVC), asbestos cement (AC), ductile iron (DI), steel and high-density polyethylene (HDPE). The studied distribution network presented 113 bursts during the observation period, between February 2014 and February 2020. 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 study. These bursts were located in 49 different pipes. Given the relatively small number of data points, the validation process will not be developed, however, in a larger dataset, this step should be carried out.

In the deterministic approach, the first step is the construction of the correlation matrix using Spearman's rank level correlation coefficient. Results are presented in Table 3 and it can be observed that the material type and year of installation are highly correlated and that the dependent variables can be grouped into two groups: Group 1 (NB, ROBLND and NBY) and Group 2 (AA10 and ROB). Variables from Group 1 presents the highest correlation values with the network properties, thus, one of these variables should be considered as dependent variable. Since the difference between the variables of Group 1 is only constant values (period of observation and average length) the choice of the dependent variable is no longer a statistically-based decision, but a technical choice. Therefore, the number of bursts per year (NBY) was chosen as the dependent variable due to the easy interpretation of the results by the water utilities.

*Table 3 – Correlation matrix using spearman's rank level correlation coefficient for Quinta do Lago – Only pipes with bursts (n=49).*

| | Mat3 | YI | DN | length | AV_T | DTT | SP | NB | AA10 | AA | ROB | ROBLND | NBY |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mat** | 1.00 | 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 | 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 | 0.19 | 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 | -0.13 | 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 | -0.11 | -0.18 | 0.03 | 1.00 | 0.02 | -0.01 | 0.06 | -0.05 | -0.11 | -0.05 | 0.06 | 0.06 |
| **DTT** | 0.07 | 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 | 0.30 | -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 | -0.51 | 0.12 | 0.34 | 0.06 | -0.22 | -0.35 | 1.00 | 0.05 | 0.63 | 0.05 | 1.00 | 1.00 |
| **AA10** | 0.06 | -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 | -0.45 | 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 | -0.09 | -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 | -0.51 | 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 | -0.51 | 0.12 | 0.34 | 0.06 | -0.22 | -0.35 | 1.00 | 0.05 | 0.63 | 0.05 | 1.00 | 1.00 |

Notes: Mat3 – Pipe material; YI – Year of installation; DN – Nominal diameter; Length – Pipe length; AV_T – Average water temperature; DTT – Distance to the tank; SP – Static water pressure; NB – Number of bursts; AA10 – Rate of burst per 100km; AA – Average age of pipe at burst; ROB – Rate of burst; ROBLND – Rate of burst length non-dimensional; NBY – number of bursts per year.

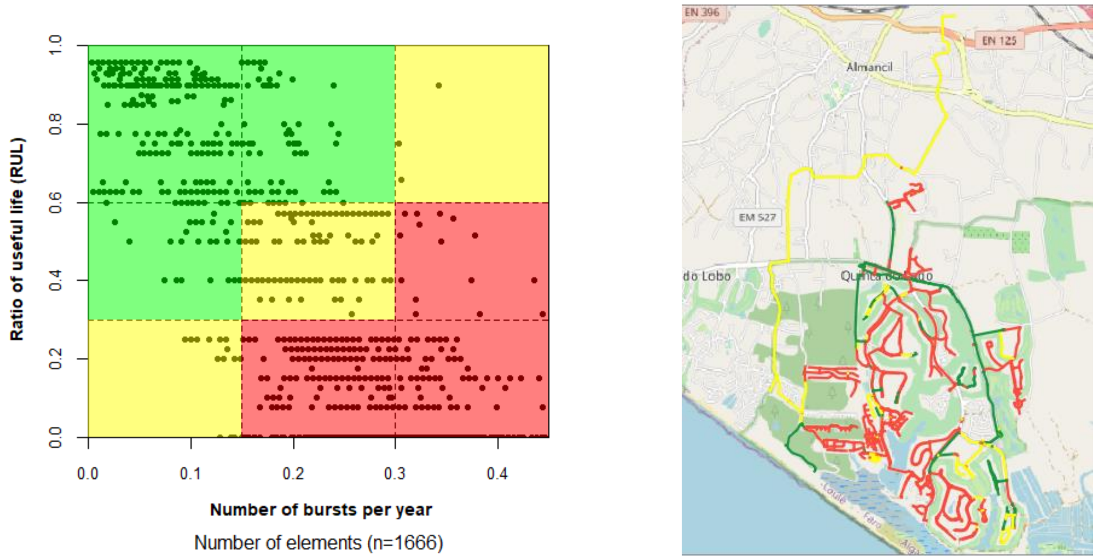
A linear regression analysis was carried out to identify the deterministic models with the best goodness-of-fit for the chosen dependent variable. Results are presented in Table 4, including the estimated regression coefficients and the coefficients of determination (i.e., $R^2$ and $R^2_{adj}$) for the best six models.

*Table 4 – Results of the 6 models with the highest coefficient of determination for the deterministic approach.*

| Model variables | $R^2$ | $R^2_{adj}$ | Y-intercept | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ |
|---|---|---|---|---|---|---|---|
| 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 |

Notes: AA – Average age of pipe at burst; SP – Static pressure; Mat3 – Pipe material; Length – Pipe length; YI – Year of installation; DTT – Distance to the tank.

The application of the linear model of average age at burst and static pressure (AA+SP) as independent variables to the case study can be observed in Figure 3(a). In which the evaluated pipes are distributed in the condition assessment matrix according to their ratio of useful life and the number of bursts per year. The good condition is associated with the green colour, the average condition with the yellow colour and the unsatisfactory condition with the red colour. The colour grid map of the physical condition is presented in Figure 3(b) allowing to identify the evaluated pipes in the distribution network. Results show that approximately 60% of the pipes are classified with an unsatisfactory condition, 28% with good condition and 12% with average condition.




*(a)* *(b)*

*Figure 3 – Results from the Deterministic approach using the linear model of the average age at burst and the static pressure: (a) condition assessment matrix; (b) colour grid map of condition.*

The heuristic approach aims to construct models that consider the perception of experts in the field to quantify the relationships between the physical deterioration of assets and their properties. A survey is developed to evaluate the influence of each factor on pipe deterioration. The developed survey divided the influencing factors into two categories, internal and external factors, and recommend experts to avoid guesswork as inaccurate results can be obtained. A total of 31 experts were surveyed and the number of experts in each experience group is similar.

The second step in the heuristic approach is the processing and analysis of the survey results to obtain quantitative weights for each influencing factor. Figure 4 presents the evaluated weights using a simple sum formulation. The vertices of the pentagon present the results of normalized weights for each expert experience group. It should be noted that the normalization of the factors must be in agreement with the already known influence of each factor on pipe deterioration.

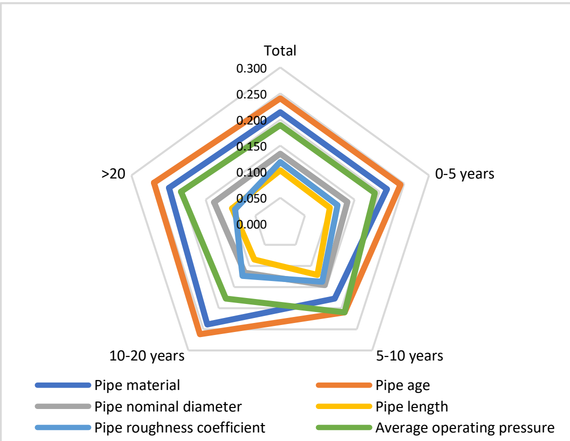
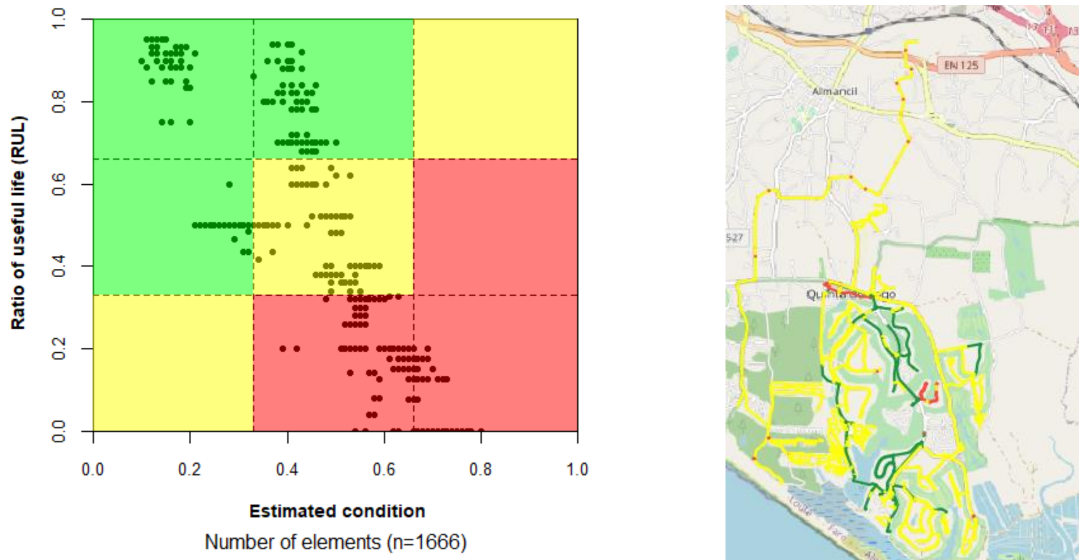



*Figure 4 – Weight associated with simple weight sum formulation.*

Results of the heuristic approach are presented in Figure 5, in which approximately 60% of the network is classified with average condition, 22% is classified with good condition and 18% is classified with unsatisfactory condition.




*(a)* *(b)*

*Figure 5 – Results from the heuristic approach using simple weight and all experts: (a) condition assessment matrix; (b) colour grid map of condition.*

A comparison between the obtained results from the two approaches (deterministic and heuristic) was carried out, allowing to demonstrate that the deterministic approach overestimates the condition classification of water distribution pipes from the case study. Assessing the performance indicator AA03, described in Alegre, Matos, *et al*., (2021), the overall pipe condition is classified as good, therefore, it is unexpected that most of the pipes are classified with average and unsatisfactory conditions through the heuristic and deterministic approaches, respectively.

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, which is that the physical condition of pipes without bursts is similar to the condition of pipes with bursts (the original dataset). Consequently, a quite penalizing approach is attained with the deterministic models.

## 4. Conclusion

The prioritization of water distribution pipes for rehabilitation requires the assessment of the physical condition of assets. A four-step methodology is proposed, including the following steps: data collection and validation, model construction using deterministic and heuristic approaches, development of a condition rating scale appropriate for each approach and, finally, condition assessment of the distribution pipes. The application of the proposed methodology allowed the assessment of the physical condition of water distribution pipes of *Quinta do Lago* in the district of the Algarve, Portugal. The obtained results from each approach presented significant differences in its 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 application of the proposed methodology a larger dataset with a higher period of observation should be carried out as future work.

## 5. Acknowledgements

The author expresses their acknowledgment to INFRAQUINTA, the water supply utility that provided the data used, as well as the *Fundação para a Ciência e Tecnologia* (FCT) for the grant provided through the project WISDOM/CEMAT.

## 6. References

Al-Barqawi, H. and 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.

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

Alegre, H. and Covas, D. I. C. (2010) *Gestão patrimonial de infra-estruturas de abastecimento de água Uma* (E. R. dos S. de Á. e R. L. N. de E. C. I. S. Técnico, ed.), Lisbon.

Alegre, H., Matos, R., Beja Neves, E., Cardoso, A., and Duarte, P. (2021) *Guia de Avaliação da Qualidade dos Serviços de Águas e Resíduos Prestados aos Utilizadores* (Instituto Regulador de Águas e Resíduos, ed.), Lisbon, Europress, Lda. [online] http://www.ersar.pt/pt/publicacoes/publicacoes-tecnicas/guias.

Alegre, H. and T.Coelho, S. (2012) Infrastructure Asset Management of Urban Water Systems. Water Supply System Analysis - Selected Topics.

Cawley, G. C. and Talbot, N. L. C. (2010) On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. , 11, 2079–2107.

Covas, D., Cabral, M., Pinheiro, A., Marchionni, V., Antunes, S., Lopes, N., Mamouros, L., and Brôco, N. (2018) Custos de construção de infraestruturas associadas ao ciclo urbano da água. , 22, 1–431. [online] http://www.ersar.pt/pt/publicacoes/publicacoes-tecnicas/guias.

Deborah J Rumsey (2016) How to Determine a p -Value When Testing a Null Hypothesis. Statistics Workbook For Dummies, Statistics II For Dummies, and Probability For Dummies. [online] https://www.dummies.com/education/math/statistics/how-to-determine-a-p-value-when-testing-a-null-hypothesis/ (Accessed April 4, 2021).

Field, R. I., Murray, D. J., Royer, M. D., and Tafuri, A. N. (2007) *Innovation and Research for Water Infrastructure for the 21St Century Research Plan*, [online] https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100ECPB.txt.

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.

Harvey, R., de Lange, M., McBean, E., Trenouth, W., Singh, A., and 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. [online] http://dx.doi.org/10.1080/07011784.2016.1224684.

King, A. P. and Eckersley, R. J. (2019) "Descriptive Statistics II: Bivariate and Multivariate Statistics" in Statistics for Biomedical Engineers and Scientists., 23–56.

Liu, Z. and 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. [online] http://dx.doi.org/10.1016/j.measurement.2012.05.032.

Neter, J., J.Nachtsheim, C., and Kutner, M. H. (1983) *Applied Linear Regression Models*, New York, NY, McGraww-Hill Irwin.

Rajani, B. and Kleiner, Y. (2001) Comprehensive review of structural deterioration of water mains: statistical models. Urban Water, 3(3), 151–164.

Royer, M. (2012) Condition assessment technologies for water transmission and distribution systems. United States Environmental Protection Agency (EPA), (March).

Selvakumar, A., Morisson, R., Sangster, T., Boyed Downey, D., and C.Matthews, J. (2013) *State of Technology for Rehabilitation of Water Distribution Systems*, Edison, New Jersy.

Thomson, J. and Royer, M. (2013) Primer on Condition Curves for Water Mains. , (December).

Ugarelli, R. and Bruaset, S. (2013) Review of deterioration modelling approaches for ageing infrastructure. SINTEF Rapport, 53(9), 1689–1699.

Urrea-Mallebrera, M., Altarejos-García, L., García-Bermejo, J., and Collado-López, B. (2019) Condition assessment of water infrastructures: Application to Segura River Basin (Spain). Water (Switzerland), 11(6), 1–21.