Pipe Life Prognosis in Water Distribution Networks using Reliable Data-based Approaches*

David Henry[1], Congcong Sun[2], Joan Vendrell[1], Vicenç Puig[1] and Enric Bonet[3]

Abstract— **The assessment and prognosis of pipe life in water distribution networks has great potential in optimizing asset investment and protecting water resources. In the state-of-the-art, most of the research work about pipe life assessment focuses on revealing associated variables and regulations for the occurrence of pipe failures, which has scientific value but still far from assisting water industry directly in real operation. In order to provide a pipe life assessment and prognosis approach with practical significance, this paper presents: 1) a comparable approach to quantify impact of different factors (mainly age, material and diameter) on the occurrence of pipe failures using statistical reliability model based on cumulative Weibull distribution, survival model based on neural networks and evolutionary polynomial regression model for pipe deterioration; 2) a prognosis method for the remaining useful life of pipes using previous algorithms; 3) a maintenance and renewal plan of the network to assist daily operation of water operators by means of a checklist including risk levels (low, medium, high) under different factor ranges. The Barcelona water distribution network is used as a real life case study, demonstrating how the proposed approaches can be used.**

I. Introduction

Pipe life assessment is a complex task which requires knowledge of the influencing variables and their association with occurrence of the pipe failures. As presented by [1], there are various investigations which collect state-of-the-art approaches on this problem from the scientific perspective and have been classified into three categories: reliability models, degradation models and data-driven statistical models. The reliability models are usually developed through considering a probability model of pipe failure that allows determining the reliability of the pipelines, as defined in [2]. The probability model depends on a parameterized function and usually considers some important factors (for instance: material, age, etc.) into the function. Different reliability models can be obtained from [3], [4] and [5], among others.

The degradation models attempt to characterize the aging process of the pipes based on their physical principles. [6] provides a comprehensive review of the physical processes of pipe degradation and classifies the models according to the following three dimensions: the smallest described element, the type of events and the modelled process. The smallest element dimension relates to whether the object of study is a pipe (pipe model) or a network of pipes (network models). Regarding the type of events, the occurrence of the failures (failure model) or the remaining useful life (RUL) is characterized, while the model tries to represent different physical degradation processes. Any combination of these three dimensions is possible, although from the available literature, only a part of these combinations have been considered. A review of statistical methods based on experimental data has been presented in [7]. In [8], evolutionary polynomial regression models are proposed to characterize the dependence of the failure rate with different feature variables (age, diameter, total length, number of pipes and failures). In [9], ANFIS-type [1] neural networks are proposed to fit an experimental regression model that relates the failure rate with diameter, age, pipe length, installation depth and pipe pressure. The results obtained with the neural network are compared with a non-linear regression, as well as with different reliability models that consider different probability distributions (Poisson and Khomsi formula, etc.). According to the results presented, the method based on neural networks works better than the rest of the methods considered. In most of these research works, the associated variables and the occurrence law of pipe failures in the WDN have been gaining more and more scientific significance [10]. However, there are still obstacles to assist the water operators in decision making on their daily management. Therefore, proposed approaches to prognosis pipe life in WDN from both scientific and practical perspective are still needed.

The main contribution of this paper includes: 1) Propose a statistical reliability model, a survival model to quantify comparably impact of different features on the occurrence of the pipe failures and an evolutionary polynomial regression (EPR) model for pipe deterioration. After considering the state-of-the-art and correlation analysis, age, material and diameter are selected as the most influencing variables to be used to characterize the pipe reliability. To describe probability distribution of the reliability model, cumulative Weibull distribution (CDF) is used. Besides, the survival model consists of Cox nonlinear proportional hazards regression with neural networks (CoxMLP). Eventually, the EPR is a hybrid regression method. 2) Pipe life evolution is used for prognosing the remaining useful life of pipes using the

[1]D. Henry, J. Vendrell and V. Puig are with Advanced Control Systems Group at the Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Llorens i Artigas, 4-6, 08028 Barcelona, Spain david.henry,vicenc.puig@upc.edu

[2]C. Sun is with Farm Technology Group at Wageningen University (WUR), P.O. Box 16, 6700 AA, Wageningen, The Netherlands congcong.sun@wur.nl

[3]E. Bonet is with CETaqua, Water Technology Centre (SUEZ Spain Group), Barcelona, 08904, Spain enric.bonet@cetaqua.com

[1]ANFIS: Adaptive Neuro-Fuzzy Inference Systems

proposed methods. 3) In order to assist water manager on their daily operation, a maintenance and renewal plan in form of easy-reading checklist is generated which includes risk levels (low, medium, high) under different physical factor ranges. All of these approaches have been demonstrated using the Barcelona WDN.

The rest of this paper is organized as follows: In Section 2, the case study is described firstly to prepare the valid data set. Besides, a preliminary analysis is also provided in order to achieve importance of different influencing factors through a pairwise correlation. In Section 3, the proposed approaches including the reliability model based on CDF, the survival model based on CoxMLP, and the EPR model are presented. After that, Section 4 includes the application results of the proposed approaches using the Barcelona WDN. Finally, in Section 5, a conclusion and future work are presented.

II. Case Study

A. Description of the available data

The case study is based on the Barcelona WDN, where historical data measurements for the pipe facilities are provided comprising 15 concerned parameters in a total number of 153285 records. The facilities were installed from the year of 1900 to 2017, while the recorded failures were recorded from 2002 to 2015. A valid data set is firstly extracted from the raw measurements following three cases: 1) The raw data includes all kinds of operations over a pipe, we only focus on the records corresponding to pipe failures; 2) We deleted the records which are lack of data. For instance, some records have 'NaN' values for pressure measurements; 3) We filter only for the information we needed. For example, there were some features related to the operation company which are not relevant to our problem.

Once the data was filtered, we compute the historical *failure rate* for each pipe. This is important because models will be developed to learn this rate from parameters of the pipe. There are different ways of defining *failure rate* as discussed in [9]. In our case study, it is defined as:

$$R = \frac{number\ of\ failures}{pipe's\ age \cdot pipe's\ length} \tag{1}$$

In other references, *failure rate* is computed using the kilometer where the leak appears. In our case study, we have considered this option as we concentrate on failures for each pipe. Note that the same pipe may have had more than one failure. Besides, the age has been considered as the difference between the *Installation Year* of the pipe and the *Failure Year* when the break was detected and the pipe was repaired.

Finally, we have worked with a data set with useful information of dimensions 153285×8, where the structure of this data set is described in the Table I, where the *Case Study* column represents the input variables in our case, the columns of *Reference No.1* and *Reference No.2* represent input variables considered in [9] and [11], respectively.

Case Study	***Reference No.1***	***Reference No.2***
Age	Age	NOPB
Material	Diameter	Material
Diameter	Length	Diameter
Length	Pressure	Length
Usage	Height	Traffic
Pressure		
Temperature		
Pressure floor's code		

TABLE I

INPUT VARIABLES ON DIFFERENT CASES.

With regards to Table I, *Pressure floor's code*, is a codification of the zone where the pipe is installed. It is very common in WDNs to divide the system into segments based on the pressure of the installation of the area. This factor can also be called as *height (or head)* in other papers like [9]. Another factor that must be explained is *Number of Previous Breaks* (*NOPB*). This input shows a different way of working with the data. In our case, we have grouped all the breaks in a pipe together. However, it could also be interesting not to do so and work with *NOPB* because once the pipe failure is repaired, the pipe has a different resistance than before. Furthermore, it is also important to explain that *Temperature* has been extracted from the historical meteorological open database of Barcelona local government. So, even this work tries to give a basic pattern to define a predictive model over WDN, the truth is that it might have some modifications depending on the initial considerations over the problem.

B. Preliminary Analysis

Generally, the more features/factors input, the better the model can be. However, more variables also means more measurements to be required, which can increase uncertainties and difficulties for real application. In order to develop an approach which can have a wide application, a preliminary analysis is taken firstly to get importance of different influencing factors. Afterwards, only the most important factors for failure rate are considered.

The preliminary analysis is based on the correlation analysis (implemented using the *pandas.DataFrame.corr* method in Python environment) using the valid data set and pipe failure rates computed in the previous section. Figure 1 shows the analysis output, which includes importance in percentage of the different considered factors. From Figure 1, it is clear that *Age*, *Material*, *Diameter* are the top three important factors to explain pipe failures since the total importance arrives 95.64%. So that, the following approaches will mainly consider *Age*, *Material*, *Diameter* as the input features. Besides, the *CodiPis* or *pressure floor* can be used to segment the whole network due to company's requirement.

III. Proposed Approaches

Three approaches, a statistical reliability model based on CDF, a survival model based on CoxMLP and an evolution-

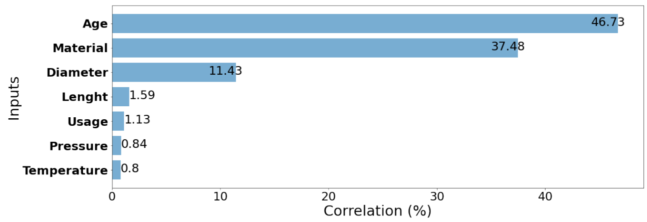

Fig. 1. Importance of different factors for failure rate.

ary polynomial regression model have been selected according to their satisfying performances in modelling numerous failure characteristics (as in [12]) and in prediction and optimization (as in [13]), respectively. Another aim of proposing two approaches is for obtaining reliable conclusions through results comparison.

A. Statistical reliability model

The statistical reliability model is obtained through learning the model parameters from historical data. This model allows to characterize the failure probability of a certain pipe at a given date.

In this way, we will have the cumulative failure probability that will allow to know for a given pipe with a certain age what is its failure probability with respect to the total number of pipes that shared that age.

To represent the analyzed data, different distribution functions may be selected in order to have a good fitting between the data and the model. According to the physical evolution of occurrence of the pipe failures, CDF as explained in [14] will be used.

The CDF that describes frequency and distribution of failures is presented in equation 2 through parameters c: the scale or characteristic life parameter; k: the shape parameter or Weibull slope and x: the considered pipe feature.

$$f(x) = \frac{k}{c}(\frac{x}{c})^{k-1}e^{-(\frac{x}{c})^k}, (x > 0, k > 0, c > 1) \quad (2)$$

To obtain the failure probability, $f(x)$ of a given material, the total number of failures in each date since the pipe has been installed is taken into account. And the failure probability of a pipe in a given date is calculated using

$$F(x) = \int_0^x f(x)\, dx = 1 - e^{-(\frac{x}{c})^k} \quad (3)$$

where c and k are the parameters that should be calibrated to achieve an optimal fitting with historical data and considering a continuous growth function tending asymptotically to 1.

B. Survival model with neural networks

After evaluating several different survival models, CoxMLP is selected due to the specific characteristics of the DWNs and their integration with neural networks providing good performance in prediction as presented in [13].

As discussed in [15], the CoxMLP approach provides a semi-parametric specification of the hazard rate, in which the instantaneous risk of failure function is expressed as a function of time and influencing factors

$$\lambda(t, X_1, ..., X_n) = \lambda_0(t)e^{\sum_{i=1}^{i=n}\beta_i X_i} \quad (4)$$

where, $\lambda_0(t)$ is the initial value of base risk. The exponential term is a relative risk that only depends on the studied covariates varying over time. X_i is the vector with multiplicative factors β_i.

This model establishes a framework to fit proportional Cox models with neural networks. From the vector of parameters, β_i of the linear predictor $g(x) = \beta_i X_i$, the neural networks have been used to parameterize $g(x)$. This parameterization of the relative risk function with a neural network does not affect the proportionality constraint of the model. Furthermore, a parametric approach without requiring stratification (grouping matching features) is proposed. In the new semi-parametric form of the Cox model, the relative risk function $h(t|x)$ depends on time as follows

$$h(t|x) = h_0(t)e^{g(t,x)} \quad (5)$$

In real application, we allow $g(t, x)$ to use time as a regular covariate, permitting $g(t, x)$ to have model interactions between time and other covariates. This is similar to the classical approach carried out in survival analysis, in which the non-proportional effect of a covariate x can be modeled by including time-dependent covariates.

Moreover, the aforementioned model is no longer a proportional hazards model, and it is still a relative risk model with the same partial probability as previously, but only considering a covariate.

C. Evolutionary polynomial regression model

The EPR is a data-driven technique based on evolutionary computations. This method deals with pseudo-polynomial structures that can represent in a precise way a physical system. This technique has several steps. First of all, EPR uses an evolutionary procedure based on a genetic algorithm (GA) to find suitable model structures. Later a linear regression step is carried and out, to compute the model constants through a least squares optimization task.

Regarding the EPR model for each material, the work proposed in [16] and [8] have been very useful to build the model of our water network (further mathematical details about EPR can be consulted in the previous papers). In [8], the model coefficient is assumed to be the same as the one given by EPR, a_1, for pipes without known failures whereas, for pipes which had failures in the monitoring period, model coefficient a_i is computed using their own burst history.

However, in our approach for every material we estimate the parameter a_k that characterizes jointly both healthy and faulty pipes by means of linear regression. When carrying out normalization step taking into account the total number of pipes, we are considering those that have not had any faults so far. This way we can model the failure rate over time independently of the existence of failures. This is that the number of recorded faults is greater than or equal to 0, $Brp \geq 0$. The steps to obtain the parameter a_k for each material are shown as follows:

1) Compute the cumulative number of failures for each age of failure in each year of the monitoring period. This value is normalized taking into account the total number of pipes and length of each material.
2) For the same age of failure in each year of monitoring, the parameter a_i is computed.
3) The a_k parameter of each material is obtained by weighing each a_i taking into account the number of failures at each age over the total number of failures.

To do this, the failure rate of pipe i according to EPR model, $\lambda_i^{EPR}(t)$, in equation 6 has been rearranged to obtain the slope, b_1, and the y-intercept, b_0, of each pipe in equation 8. Note that b_1 returns the unknown a_i.

$$\lambda_i^{EPR}(t) = \frac{a_1}{T} \frac{Lp_i(A_{0,class} + t)}{D_{class}^{1.5}} \tag{6}$$

where T is the monitoring period, Lp_i is the length of the i_{th} pipe, $A_{0,class}$ is the equivalent age of the pipe class at the end of the monitoring period, t is the time variable and D_{class} is the equivalent diameter of the pipe class. D_{class} can be obtained in equation 7.

$$D_{class} = \frac{\sum_{class}(L_p \cdot D_p)}{L_{class}} \tag{7}$$

$$\lambda_i^{EPR}(t) \cdot \frac{T \cdot D_{class}^{1.5}}{Lp_i} = a_1 \cdot t + a_1 \cdot A_{0,class} = b_1 \cdot t + b_0 \tag{8}$$

Thus, the failure rate of each pipe can be obtained in equation 9. Moreover, the number of leakages predicted can be obtained by means of the integration for all pipes in equation 10.

$$\lambda_i(t) = \frac{a_k}{T} \frac{Lp_i(A_{0,class} + t)}{D_{class}^{1.5}}, \quad \text{if } Brp_i \geq 0 \tag{9}$$

Moreover, the number of leakages predicted for each pipe, BR_i, can be obtained by means of the integration for all pipes in equation 10 considering the time horizon, h.

$$BR_i = \int_0^h \lambda_i(t) \cdot dt \tag{10}$$

Eventually, the failure probability can be computed in equation 11.

$$Failure\ probability\ (\%) = \frac{\sum BR_i}{\sum pipes} \tag{11}$$

D. Prediction and prognosis

As it has been aforementioned, CoxMLP is able to predict failure probability according to the desired features. Moreover, we can also implement CoxMLP to pipe prognosis using the non-linear neural network framework as discussed in [17], as well as removing to a bigger extent the proportionality constant of the Cox model. For this purpose, a 25 years ahead prediction in steps of 5 years has been carried out while taking into account only the most important features: age, material and diameter. In addition, Weibull distributions and the built EPR model will be used to predict failure rate in the aforementioned time dates.

There are in total 18 kinds of material used in the Barcelona WDN, the main ones are: ductile iron (41.69%), gray cast iron (22.35%), asbestos cement (13.92%), high density polyethylene (10.88%), low density polyethylene (5.60%), reinforced concrete (1.50%) and reinforced concrete welded joint (1.48%). The usage of different materials takes different percentage (as shown after each material) of the whole network. The material with less usage does not have enough data. So that in the prognosis, we focus on the material with more usage. Thus, we will carry out failure probability prediction only for asbestos cement, gray cast iron and ductile iron. Thus, altogether with material, the features of age and diameter will be taken into account as well in the different forecasts.

IV. Results

Results for statistical reliability model, survival model, evolutionary polynomial regression model, failure rate prognosis and probability prediction will be provided in this section.

A. Failure model validation

The Barcelona WDN incorporates several kinds of pipes according to their materials used in the construction process and period. The representative materials (mainly asbestos cement, gray cast iron and ductile iron) allow us to develop the reliable and survival models in terms of RUL over time. Figure 2 presents the RUL evolutions for the material feature in Gray cast iron in 118 years using reliable model based on CDF and the survival model based on CoxMLP, respectively. For a better interpretation of the figures the vertical axis contains the RUL probability, that goes from the probability value "0" meaning that no failures have been experienced and that there will be less remaining useful life as the years go by. Taking a closer look, it is shown that the RUL starts to decrease sharply from 0 at beginning. From around 50 to around 100 years, the RUL becomes smoothly, and then increase sharply afterwards tending asymptotically to 1.

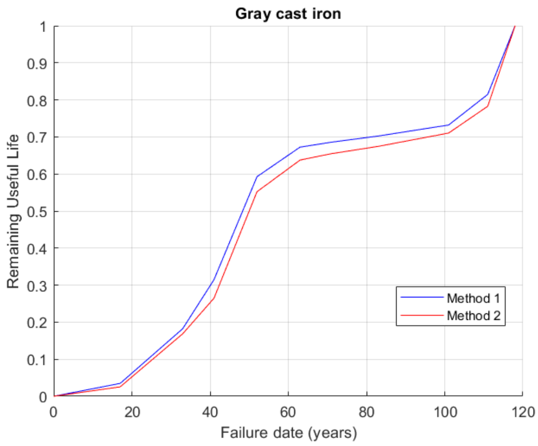

Fig. 2. RUL based on Weibull and CoxMLP.

To have a clear comparison for the performance between the two approaches, Table II is provided showing more details. Note that this analysis has been performed for a failure subset within the entire water network. So, the conclusions will be limited to the previous subset not to the entire set.

RUL analysis for materials (failure probability in %)					
Material	$n^{\underline{o}}_{\text{failures}}$	Date (years)	1st method	2nd method	Difference (%)
Gray cast iron	4258	(63-101)	67.2 %	63.8 %	3.4 %
Ductile iron	2985	(59-102)	71.4 %	67.8 %	3.6 %
Asbestos cement	1844	(56-101)	74.9 %	70.8 %	4.1 %

TABLE II

RUL DIFFERENCE COMPARISON REGARDING FAILURE PROBABILITY.

From this table, it should be clarified that, the materials which have a higher usage percentage in the whole network do have more failures after we take into account the age and material features using CDF method (1st method), as shown in the column $n^{\underline{o}}_{\text{failures}}$.

However, the CoxMLP method (2nd method) includes more parameters that intervene in the failures, hence, the percentage of RUL is lower, due to the fact that each intervening factor will have a given weight on the total figure. Some of the factors that were not considered in method 1, embracing the maximum and the minimum average pressures due to the lack of data, will decrease the factual percentage.

As the number of failures decreases for the less used materials, the second method tells us that the RUL is lower and, therefore, that a higher percentage of failures has occurred according to a lower relative weight of material respect to the other factors that intervene in failures.

The fact of including a greater number of factors in the second method will allow us to obtain a RUL curve that represents more precisely the pipes behaviour for each studied material. In order to characterize this behaviour, the number of input data incorporated into the model must be sufficiently representative so that the incidence of each variable suits better to reality.

B. *Failure rate prognosis*

In order to apply the results into reality with application wise to facilitate the labour of the WDN operators, a pipeline monitoring and maintenance plan has been developed regarding the most important factors (age, material and diameter) in form of an easy-reading checklist. Four assessment zones (A, B, C, D) and three operations (continue to work, should be supervised, should be replaced) are considered following this procedure:

1) The assessment starts from zones B and C, which corresponds to the [65 - 70]% and [75 - 80]% of the age, where the failures in low appearance starts and ends, respectively.

2) The two zones A and D correspond to the periods before and after the assessment stages. In zone A, the failure frequency will be very low. Zone D represents the time when many failures will happen. So that, it must be ensured that the asset will be replaced before arriving zone D where the probability of failure is very high.

3) Zones B and C have been used as thresholds to assess whether the pipes can continue to work, should be supervised or should be replaced immediately, taking into account the parameters which have a greater impact on occurrence of failures. More details about the usage of Zone B and Zone C are provided as follows:

Zone B:

- Only one parameter is out of the range (values higher than the threshold). Active entering the monitoring phase.
- More than one parameter is out of the range. Asset replacement is considered.

Zone C:

- Only one parameter is out of the range. Asset replacement is considered.
- More than one parameter is out of the range. The replacement action of the asset is activated.

C. *Failure probability prediction*

The failure probability $FP(\%)$ will be obtained taking into account the number of pipe leakages divided by the total number of pipes. As we are carrying out a prediction, new leakages will be added to the total number of leakages in a cumulative manner. The failure probability prediction result focusing on the three main materials in terms of each 5 future years for the 3 developed methods is provided in Table III.

For the EPR method, the parameter a_k obtained for each material can be consulted as follows:

- Gray cast iron $a_k = 0.0858$

- Ductile iron $a_k = 0.0199$
- Asbestos cement $a_k = 0.0560$

FAILURE PROBABILITY PROGNOSIS IN THE FUTURE 25 YEARS				
		EPR	Weibull	COX
Material	Time horizon	FP (%)	FP (%)	FP (%)
Gray cast iron	5 years	0.8589	0.9214	0.7921
Gray cast iron	10 years	1.7789	2.0871	1.9762
Gray cast iron	15 years	2.7601	2.9584	2.9042
Gray cast iron	20 years	3.8023	3.8848	3.5587
Gray cast iron	25 years	4.9057	4.9879	4.9247
Ductile iron	5 years	0.2540	0.3256	0.2696
Ductile iron	10 years	0.5429	0.7393	0.6741
Ductile iron	15 years	0.8668	1.0503	0.9929
Ductile iron	20 years	1.2258	1.3822	1.2193
Ductile iron	25 years	1.6197	1.7781	1.6395
Asbestos cement	5 years	0.8994	0.9537	0.8749
Asbestos cement	10 years	1.8556	2.1618	2.0638
Asbestos cement	15 years	2.8685	3.0661	2.9728
Asbestos cement	20 years	3.9382	4.0284	3.8419
Asbestos cement	25 years	5.0647	5.1749	5.0454

TABLE III

K-YEAR AHEAD FAILURE PROBABILITY PREDICTION.

It can be seen that all predictions given by the 3 methods are similar among them. The probabilities obtained by EPR and COX are very similar to each other, while in the Weibull case the probabilities are slightly higher.

V. Conclusions

This paper proposes a comparable approach to quantify impact of different factors (mainly age, material and diameter) on occurrence of pipe failures using statistical reliability model based on CDF, survival model based on CoxMLP and an evolutionary polynomial regression (EPR) model for pipe deterioration. The produced RUL evolution in 118 years confirm similarities (less than 5% difference) of both approaches. The CoxMLP survival model has lower RUL and suits better in reality due to more involvement of factors. Besides, more conclusions can also be obtained that the materials with higher percentage usage do have more failures. For application wise, the prognosis checklist in terms of different assessment zones (A, B, C, D) with corresponding operations (continue to work, should be supervised, should be replaced) is also provided. For more clarification, a 25-year failure probability prediction is also added in an easy-reading table to evaluate levels of risk (Low risk zone, Medium risk zone and High risk zone) for each pipe in the coming years with significance in assets monitoring and maintenance from both scientific and practical perspectives.

Acknowledgements

This work has been funded by the Catalan Agency for Management of University and Research Grants (AGAUR) ACCIO RIS3CAT UTILITIES 4.0 P4 ACTIV 4.0 and Grant 001-P-001643 Agrupació Looming Factory. Financial support also from NWO (Netherlands Organisation for Scientific Research) Sectorplan for Beta and Technology Programme in Wageningen University.

References

[1] D. Winkler, M. Haltmeier, M. Kleidorfer, W. Rauch, and F. Tscheikner-Gratl, "Pipe failure modelling for water distribution networks using boosted decision trees," *Structure and Infrastructure Engineering*, vol. 14, no. 10, pp. 1402–1411, 2018.

[2] K. Pietrucha-Urbanik and K. Pociask, "Analysis and assessment of water distribution subsystem failure," *Journal of KONBiN*, vol. 40, pp. 47–62, 2016.

[3] Y. C. Su, L. W. Mays, N. Duan, and K. E. Lansey, "Reliability based optimization model for water distribution systems," *Journal of Hydraulic Engineering*, vol. 114, no. 12, pp. 1539–1556, 1987.

[4] D. Khomsi, G. A. Walters, A. R. D. Thorley, and D. Ouazar, "Reliability tester for water distribution networks," *Journal of Computing in Civil Engineering*, vol. 10, no. 1, pp. 10–19, 1996.

[5] O. Fujiwara and H. Tung, "Reliability improvement for water distribution networks through increasing pipe size," *Water Resource Research*, vol. 27, no. 7, pp. 1395–1042, 1991.

[6] Y. Kleiner and B. Rajani, "Comprehensive review of structural deterioration of water mains: statistical models," *Urban Water*, vol. 3, no. 3, pp. 131–150, 2001.

[7] A. Scheidegger, J. P. Leitão, and L. Scholten, "Statistical failure models for water distribution pipes – a review from a unified perspective," *Water Research*, vol. 83, pp. 237–247, 2015.

[8] L. Berardi, O. Giustolisi, Z. Kapelan, and D. A. Savic, "Development of pipe deterioration models for water distribution systems using epr," *Journal of Hydroinformatics*, vol. 10, no. 2, pp. 113–126, 2008.

[9] M. Tabesh, J. Soltani, R. Farmani, and D. Savic, "Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling," *Journal of Hydroinformatics*, vol. 11, no. 1, pp. 1–17, 2009.

[10] C. C. Sun, B. Parellada, V. Puig, and G. Cembrano, "Leak localization in water distribution networks using pressure and data-driven classifier approach," *Water*, vol. 1, no. 12, p. 54, 2020.

[11] S. Christodoulou and A. Deligianni, "A neurofuzzy decision framework for the management of water distribution networks," in *Water Resour Manage*, S. Science, Ed., 2009, pp. 139–156.

[12] R. Gifty and R. Bharathi, "Weibull cumulative distribution based real-time response and performance capacity modeling of cyber–physical systems through software defined networking," *Computer Communications*, vol. 150, pp. 235–244, 2020.

[13] T. Awolusi, O. Oke, A. Sojobi, and O. Aluko, "Performance comparison of neural network training algorithms in the modeling properties of steel fiber reinforced concrete," *Heliyou*, vol. 5, p. e01115, 2018.

[14] G. Vladeanu and D. Koo, "A comparison study of water pipe failure prediction models using weibull distribution and binary logistic regression," *Water Research*, vol. 83, pp. 237–247, 2015.

[15] H. Kvamm, Borgan, and I. Scheel, "Time-to-event prediction with neural networks and cox regression," *Machine Learning Research 20*, pp. 6–8, 2019.

[16] G. Orazio and S. Dragan, "A symbolic data-driven technique based on evolutionary polynomial regression," *Journal of Hydroinformatics*, vol. 8, pp. 207–222, 2006.

[17] J. L. Katzman, U. Shaham, A. Cloninger, J. Bates, T. Jiang, and Y. Kluger, "Deepsurv: personalized treatment recommender system using a cox proportional hazards deep neural network," *BMC Medical Research Methodology*, vol. 18, no. 1, p. 24, 2018.