Applied Soft Computing 111 (2021) 107731

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

An evolutionary fuzzy system to support the replacement policy in water supply networks: The ranking of pipes according to their failure risk

Alicia Robles-Velasco [a,b,*], Jesús Muñuzuri [a], Luis Onieva [a], Pablo Cortés [a,b]

[a] Dpto. de Organización Industrial y Gestión de Empresas II. ETSI. Universidad de Sevilla, Spain
[b] Cátedra del Agua (EMASESA-Universidad de Sevilla), Spain

A R T I C L E I N F O

Article history:
Received 1 June 2020
Received in revised form 26 February 2021
Accepted 15 July 2021
Available online 24 July 2021

Keywords:
Evolutionary fuzzy system
Fuzzy logic
Genetic algorithm
Water supply networks
Pipe failure prediction

A B S T R A C T

In this study, an evolutionary fuzzy system is proposed to predict unexpected pipe failures in water supply networks. The system seeks to underpin the decisions of management companies regarding the maintenance and replacement plans of pipes. On the one hand, fuzzy logic provides high degrees of interpretability over other black box models, which is requested in engineering application where decisions have social consequences. On the other hand, the genetic algorithm helps to optimize the parameters that govern the model, specifically, for two purposes: (i) the selection of variables; and (ii) the optimization of membership functions.

Data from a real water supply network are used to evaluate the accuracy of the developed system. Several graphs that depict the ranking of pipes according to their risk of failure against the network length to be replaced support the choice of the most successful model. In fact, results demonstrate that the annual replacement of 6.75% of the network length makes it possible to prevent 41.14% of unexpected pipe failures.

1. Introduction

Water supply networks are infrastructures that transport drinking water from treatment plants to consumption points. They play a key role in the economic and social development of cities because they provide a basic resource that people and industries need on a daily basis. According to the 2019 Human Development Report published by the United Nations [1], countries with higher human development have quality and safe water supply networks. In order to maintain these quality levels, management companies must avoid security risks and supply disruptions as much as possible. Since pipes are the main components of water supply networks, one of their priorities must be to prevent pipe failures.

Controlling the state of water supply networks is not an easy task since most of the pipes are buried, which hinders access and, therefore, their maintenance. An inadequate management of these infrastructures can result in an increase of unexpected pipe failures, causing serious problems. On the one hand, large quantities of water are lost which entails a decrease in system sustainability and monetary losses. On the other hand, supply disruptions can cause water pollution due to pressure losses, as well as a security risk to the population. To tackle this problem, companies in charge of the management of these infrastructures have begun to consider machine learning as prediction tool. If a system is too complex or handles large amounts of data, expert knowledge is insufficient. On the contrary, machine learning can extract hidden patterns from large amounts of data, thus being a perfect solution to support decision-making processes.

Many studies have employed statistical models to predict pipe failures in water supply networks. Survival models are used to predict the time to failure of pipes [2–4] which is useful to analyze the global state of the network. Logistic regression allows estimating the probability of failure [5,6], helping to take measures regarding specific supply or sewer pipes [7]. Bayesian Belief Networks (BBN) combine diagnosis and predictive analysis and are used to estimate risk index per area [8,9]. In general, these statistical models have a limited learning capacity. Artificial neural networks (ANN) and support vector machines (SVM) overcome this weakness and enhance more accurate results. Several studies have employed these methodologies to predict the failure rate [10–12] the time to failure [13], or the failure probability [14] of water supply pipes. Nevertheless, both ANN and SVM are black box systems, i.e., it is almost impossible to decipher the connection among the variables involved in their functioning. As a solution, we propose an Evolutionary Fuzzy System to predict the

* Corresponding author at: Dpto. de Organización Industrial y Gestión de Empresas II. ETSI. Universidad de Sevilla, Spain.
E-mail address: arobles2@us.es (A. Robles-Velasco).

https://doi.org/10.1016/j.asoc.2021.107731
1568-4946/

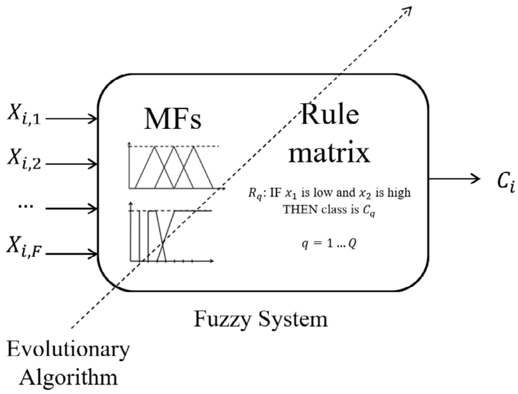

Fig. 1. Evolutionary fuzzy systems.

future pipe failures based on the historical database of a company. This system provides a simple rule-based matrix that connects the explanatory variables with the risk of failure, so the interpretability of results is assured. Moreover, the learning capacity is guaranteed thanks to the use of evolutionary algorithms.

Departments that manage supply water networks must justify and defend their decisions in front of a committee, moreover, finding specific causes of pipe failures holds great value. For example, if a range of diameters of a kind of material poses a high risk of failure, its installation should be avoided in future. Furthermore, fuzzy logic facilitates the handling of linguistic and inexact variables which are common in water supply systems. For this reason, it is an excellent option. A fuzzy logic system provides a set of easily interpretable linguistic rules which allow experts and users to understand the problem [15]. Moreover, experience has shown that this technique allows achieving levels of accuracy as high as ANN or SVM. Fuzzy logic has been previously applied to predict pipe failures in water supply networks [16–21]. In all these studies, expert opinions were used to establish the components of the fuzzy system, which does not guarantee its optimization. Evolutionary fuzzy systems (EFSs) overcome this weakness by fixing the parameters that govern fuzzy logic systems based on real data from the network. They use evolutionary algorithms (EAs), mainly genetic algorithms, to search for the optimal parameters in a solution space.

A more detailed literature review of EFSs is presented below. In Fig. 1, the arrangement of EFSs and the connection between fuzzy logic and evolutionary algorithms is schematically shown. The fuzzy logic system, which is represented inside the box, can be broadly divided into two stages or modules; the input variables are first fuzzified by the membership functions (MFs), and then the rule matrix relates these fuzzified variables with the output variable. The EAs can be implemented in these two main stages of the fuzzy logic system. Some of the most common uses are: (i) optimization of membership functions (shape, number of fuzzy sets or position of these fuzzy sets); and (ii) generation and selection of the rules that compose the rule matrix. The optimization of MFs involves many parameters. For this reason, simple heuristics are sometimes used to fix some of them, and more complex evolutionary algorithms as GA or PSO (Particle Swarm Optimization) are used to seek the optimal values of others. In [22], the shape and number of fuzzy sets are initially established, and EAs are used to vary the position of the fuzzy sets. In [23], four approaches, free and bounded with binary-coded and real-coded GA, are employed to optimize trapezoidal MFs. Their results sustained that free optimization and real-coded GA are the most suitable option for this purpose. Alcalá et *al.* [24] designed four different MFs, which have a total of fourteen fuzzy sets, and search for the optimal ones to each variable. In order to apply the same MFs to all variables, they are previously normalized (values from 0 to 1) which produces a loss of interpretability since rules are based on transformed variables. Some studies include the selection of variables as parameters of the EA and others employ dependency-based measures as the Mutual Information [25].

The optimization of the rule matrix starts with the generation of a set of rules. Then, rules are modified and selected using some EA according to certain criteria, such as support or confidence. The scalability of the problem is a critical aspect: the use of more variables supposes an exponential increase on the number of rules or the rules' length. Consequently, the training time grows considerably due to EAs' expensive computation.

Due to the wide range of possibilities related to the design of EFSs, various studies propose their own system. Table 1 contains the main characteristics of seven EFSs. As recommended by Fernandez et al. [26], all the studies verify the performance of their proposed EFSs using datasets available in public databases. Additionally, they apply other previous EFSs to the same datasets in order to compare the results. The size of the datasets of Table 1 varies from very small, less than 100 samples, or small which has less than 1000 samples, to big data applications.

The applications of these studies are also pointed in the table. Ferranti et al. [27] perform a big data application of an EFS, minimizing the accuracy while maximizing the interpretability of results. To demonstrate its performance, they apply their designed algorithm to ten datasets. Ganesh et *al.* [25] use an EFS to attempt medical applications whose most important feature is that the number of explanatory variables (genes) is much higher than the number of samples. Cózar et al. [15] develop a metaclassifier to compare the performance of several EFSs. In this study, IVTRS-Imbalanced is presented, which is an EFS system that attains small sets of fuzzy rules and, therefore, highly interpretable models [28]. It is focused on financial problems that require transparency and interpretability. It is interesting to highlight that fuzzy logic is a potential tool to solve important problems in the real-time control field [29].

Regarding the evaluation function of EAs, there are approaches with a single objective and Multi-Objective Evolutionary Algorithms (MOEA). On the one hand, single objective approaches seek to maximize the percentage of correct predictions or the accuracy. In general, the number of rules and their length is free or previously established. On the other hand, MOEA approaches also try to minimize the number of rules which compose the rule matrix or its length, measuring the number of antecedents of each rule. Some authors have noticed that the maximization of EFS accuracies usually results in the reduction of their interpretability [30]. Consequently, for cases where the importance does not only lie in making good predictions, but also in interpreting results, MOEA approaches are more convenient. Antonelli et *al.* [31] employ a MOEA to create a fuzzy system for classification tasks. They defend that their EFS, called PAES-RCS, considerably reduces the number of evaluations to achieve the same levels of accuracies that two other well-known approaches. Aghaeipoor and Javidi [32] implement a two-phase MOEA focused on high-dimensional regression data. It has two objectives: the first one is to select the more significant attributes and to tune the shape and position of membership functions, and the second one is to reduce the rule base by eliminating the weakest rules. The selection of rules is made based on their support and confidence.

In this paper, a single-objective EFS is designed for forecasting pipe failures in water supply networks. As a contribution to the field, we have developed a fuzzy model specifically adapted to the characteristics of the problem under study. For this purpose, the proposed genetic algorithm addresses two aspects: (i) the selection of variables; and (ii) the optimization of membership

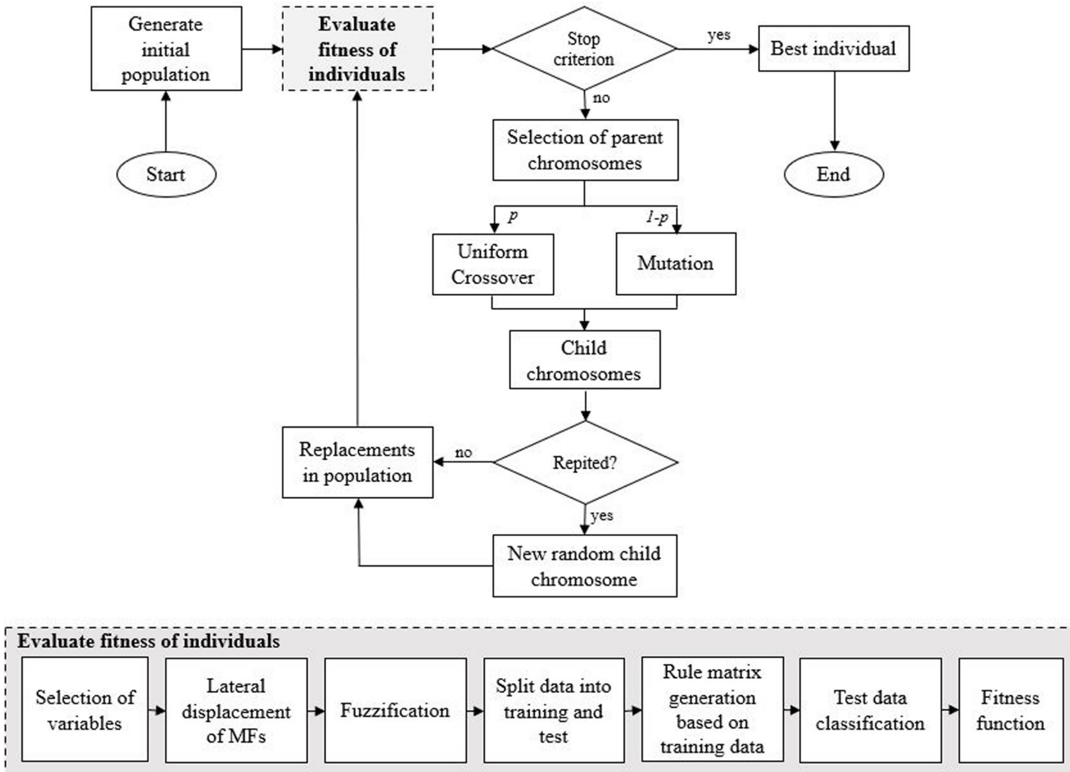

Fig. 2. Architecture of the proposed methodology.

Table 1
Evolutionary Fuzzy systems proposed in the literature and their main characteristics.

Ref.	EFS	Objective	Data base	Rule base	Applications
[24]	NSGAII-SG	Max Acc. + Min no. and length of Rules	Lateral tuning of MFs with GA	Selection of rules with GA	Medical. Only continuous data. Small datasets.
[25]	GSA	Min no. wrong classifications + Min no. rules	Lateral tuning of MFs with PSO and mutual information technique to select variables	Selection of rules with GA	Genetics. Very small datasets
[22]	AGFS	Max Acc.	Selection of variables with GA	Generation of rules with GA	Small datasets
[31]	PAES-RCS	Max Acc. + Min Rules' total length	Lateral tuning of MFs with GA and selection of variables through rules	Selection of rules with GA	Small datasets
[28]	IVTURS-Imbalanced	Max AUC	Lateral and shape tuning of MFs with GA	Selection of rules with GA	Financial applications. Imbalanced data.
[27]	DPAES-RCS	Max Acc. + Min Rules' total length	Lateral tuning of MFs with GA and selection of variables through rules	Selection of rules with GA	Big Data
[32]	MOKBL+ MOMs	Min NRMSE + Min no. and length of Rules	Lateral and shape tuning of MFs with GA	Selection of rules with GA	High-dimensional regression data

functions. This makes the system more flexible and independent, giving data the power to decide which variables and which intervals are the most discriminatory. As a novelty, we analyze the link between the risk of failure and the length of the network by means of a series of graphs. This is useful to justify the results and to evaluate the applicability of the methodology as a decision support system. To the best of our knowledge, it is the first attempt to apply this technique with this purpose, having been successfully applied to real problems in other areas. For instance, to discover the input–output relationship between the soil spectra and the soil properties [33], to predict suspended sedimentation of rivers [34], or to predict monthly river flows [35].

The paper is organized in the following sections. Section 2 presents the proposed methodology, focusing on the characteristics of the fuzzy model and the genetic algorithm. In Section 3, the implementation of this methodology to a real case study is developed, including the description and processing of data, the calibration of the system and the results. Moreover, an adjustment of the methodology is also performed. A more extensive set of results is presented in the Appendix, at the end of the manuscript. Section 4 contains a discussion about the best obtained solution and a qualitative analysis of rules. Finally, conclusions are presented in Section 5.

$$x_i \longrightarrow \boxed{\text{MFs}} \longrightarrow \mu_{A_{f,j}}(x_i) \quad \forall f = 1, \ldots, F;\ \forall j = 1, \ldots, T_f$$

Fig. 3. Fuzzification of variables.

2. Proposed methodology: evolutionary fuzzy system

The EFS designed in this study utilizes a GA for two purposes: (i) to select the most relevant input variables; and (ii) to optimize MFs through the lateral displacement of their fuzzy sets. Fig. 2 shows the main steps of the proposed methodology. The implementation of the GA consumes large runtimes because a new fuzzy system must be built based on training data to evaluate each individual of the GA. However, it allows seeking the fuzzy system with the highest classification capabilities.

Firstly, we introduce the main characteristics of the fuzzy system in Section 2.1, and secondly, in Section 2.2 we describe the components and parameters of the GA.

2.1. Fuzzy system

Fuzzy logic (FL) was established by Lotfi Zadeh in 1965 through a study named *Fuzzy Sets* [36]. While classical logic maintains that everything can be represented in binary terms, fuzzy logic uses degrees of truth. This allows the partial membership of sets and it is preferred when the interaction between variables and the system behavior is not completely understood.

The following subsections present the description of the main modules of our fuzzy system: fuzzification, rule matrix and classification.

2.1.1. Fuzzification

Fuzzification is the process of assigning to each real variable its fuzzy values (numbers from 0 to 1). This is done using membership functions. There is a great variety of MFs like triangular, trapezoidal, gaussian, etc. Fig. 3 shows the fuzzification of a sample i through a function that links the elements of its domain or universe of discourse U_f with elements of the Interval [0, 1]. In this study, triangular MFs are chosen because they have achieved good results in a wide range of problems, and it simplifies the operation of the EA.

Let N be the number of samples that constitute the datasets, and $x_i = (x_{i1}, \ldots, x_{if})\ \forall i = 1, \ldots, N$ be the samples composed of $f = 1, \ldots, F$ explanatory variables. Variables can be continuous or discrete. Continuous variables usually come from numerical measures while discrete variables typically arise from categorical real data. Continuous variables are defined in continuous universes of discourse $U_f \in R$ which have their own fuzzy sets ($A_{f,j}$). The number of fuzzy sets included in each universe of discourse is called T_f. Therefore, $j = 1, \ldots, T_f$ are the sub-index associated with the different fuzzy sets of each variable f. Eq. (1) describes a triangular fuzzy set where a, b and c are the positions of its vertices in the universe of discourse of a variable f.

$$\mu_{A_{f,j}}(x_i) = \begin{cases} 0 & if\ x \le a \\ \frac{x-a}{b-a} & if\ a < x \le b \\ \frac{c-x}{c-b} & if\ b \le x < c \\ 0 & if\ x \ge c \end{cases}$$
$$\forall f \in Numerical\ variables \quad \forall j = 1, \ldots, T_f \qquad (1)$$

The number of partitions or fuzzy sets associated to each explanatory variable has a direct relationship with the interpretability of results. The greater the number of partitions, the less interpretable they are. Nevertheless, a very small number of partitions may cause a substantial loss of predictive accuracy. In this study, the number of fuzzy sets of numerical variables varies from 3 to 5 ($T_f = 3, 4, 5$). Moreover, they are strong fuzzy sets. MFs are initially uniform and core displacements are allowed by means of the GA. Fig. 4 shows three initial MFs with 3, 4 and 5 partitions respectively.

It should be noted that the membership to fuzzy sets of numerical variables varies from 0 to 1, i.e. $\mu_{A_{f,j}}(x_i) \in [0, 1]$. The closer to 1, the greater membership of the sample x_i to the fuzzy set $A_{f,j}$. Discrete variables are defined in finite or discrete universe of discourses U_f. In this case, T_f is the number of categories of the variable f, and the membership of samples to fuzzy sets is either complete ($\mu_{A_{f,j}}(x_i) = 1$) or null ($\mu_{A_{f,j}}(x_i) = 0$).

2.1.2. Rule matrix

Rule matrix is the inference system between inputs and outputs. Each rule is composed of a set of antecedents and a consequent. Antecedents are the conditions that must be satisfied so that consequent occurs, in this case, the assignment of a class to a sample. Two well-known fuzzy logic-based systems are Mamdani [37] and Takagi–Sugeno [38]. The main difference between them is the format of rule consequents. The rule consequent of Mamdani models is a linguistic expression. Whereas in Takagi–Sugeno models, consequents are calculated through linear functions that connect input variables. In this study, the Mamdani-type Inference System is chosen because it allows working with linguistic variables and it achieves higher levels of interpretability. Moreover, these systems have been used in all the revised EFS. Rules have the following structure:

$$R_q\text{: } If\ x_{i,1}\ is\ A_{q,1_j}\ and \ldots and\ x_{i,f}\ is\ A_{q,f_j}\ then\ y_i\ is\ C_q\ with\ RW_q$$

Where Q is the set of fuzzy rules ($q = 1, \ldots, Q$) that composes the rule matrix. $A_{q,f}$ represents the fuzzy set of variable f that is antecedent of rule q. As previously said, variables can have different numbers of fuzzy sets T_f, therefore, $A_{q,f} = A_{f,j}$ with $j = 1, \ldots, T_f$. Let us denote by $C_q \in \{0, 1\}$ the class or consequent of rule q. C_q takes the value of 1 if the rule predicts that pipes will break, and 0 otherwise. Finally, RW_q is the weight of the rule q within the rule matrix. This indicates its importance or relevance making classifications. Both $x_i = (x_{i,1}, \ldots, x_{i,f})$ and y_i are respectively the inputs and the output corresponding to sample i.

To generate a rule, we must first generate its antecedents. A rule is activated if all its antecedents are fulfilled, and we have forced to include at least one fuzzy set of each variable in each rule. Furthermore, as many rules as combinations of fuzzy sets of the different variables are generated, ensuring the system to be complete, which means that for any sample the system activates at least one rule. Consequently, an output is always assigned to all samples.

Secondly, a class or consequent must be assigned to each rule, which is a more complex task. In traditional fuzzy systems, the consequent of the rules is chosen by experts. However, this is not feasible with large number of rules. Besides, expert opinions are subjective and can vary from a dataset to another. For this reason, in this study, both rule classes C_q and rule weights RW_q are established based on historical data. The criterion followed by [24] is adopted. First, we calculate the matching degree of each training sample x_i with the antecedents of each fuzzy rule R_q using the product operation as shown in Eq. (2). Where $\mu_{A_{q,f_j}}(x_i)$ is the membership of the sample i to the antecedent fuzzy set $A_{f,j}$ present in rule q.

$$w_q(x_i) = \prod_{f=1}^{F} \mu_{A_{q,f_j}}(x_i)$$

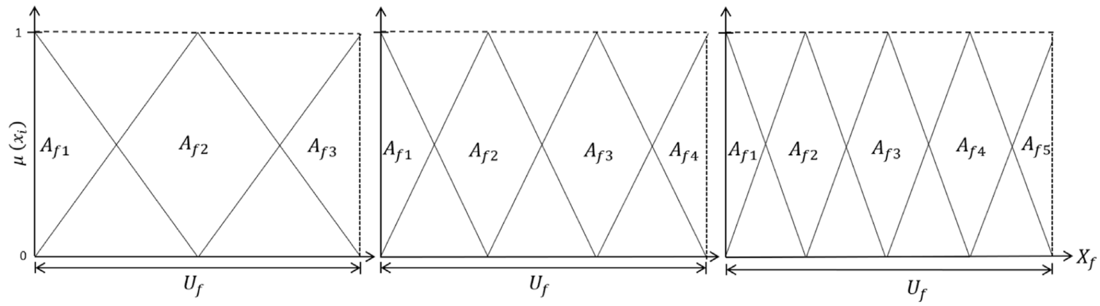

Fig. 4. Triangular and strong membership functions of numerical variables with 3, 4 and 5 partitions or fuzzy sets.

$$\forall i = 1, \ldots, N;\ \forall q = 1, \ldots, Q \tag{2}$$

Then, the class with a major confidence with the rule (Eq. (3)) is assigned to each rule. The confidence of a rule with a class is the sum of the matching degrees with samples of this class among the sum of the matching degree with all samples (of both classes). This is done using the training data.

$$c\left(R_q \to C_q\right) = \frac{\sum_{x_i \in claseC} w_q\left(x_i\right)}{\sum_i w_q\left(x_i\right)} \quad \forall q = 1, \ldots, Q \tag{3}$$

The rule weights are computed as the difference between the confidence of one rule with its class and its confidence with the opposite class (Eq. (4)).

$$RW_q = c\left(R_q \to C_q\right) - c\left(R_q \to Class \neq C_q\right) \quad \forall\ q = 1, \ldots, Q \tag{4}$$

The support of a rule (Eq. (5)) measures the total matching degree of the samples with the rule among the total number of samples. To this purpose, it only includes the matching degree of samples with the same class than the rule. The higher the support, the higher is the coverage of the rule. It is an interesting metric to detect the most general rules. For instance, a rule which only covers one sample will have a confidence of 1, being a very specific rule. This rule is less significant than others with higher supports but slightly lower confidences. Support is employed to do the post-analysis of rules in Section 4.

$$s\left(R_q \to C_q\right) = \frac{\sum_{x_i \in claseC} w_q\left(x_i\right)}{N} \quad \forall\ q = 1, \ldots, Q \tag{5}$$

The total number of rules (TNR) depends on the number of variables and the number of fuzzy sets. Being T_f the number of fuzzy sets of continuous variables and T_k the number of categories of categorical variables, TNR is calculated by Eq. (6). Since it is a general equation, the number of continuous variables has been represented as V_{cont} and the number of categorical ones as V_{cat}. Additionally, it has been assumed that all categorical variables have the same number of categories. If this was not the case, the product would include as many factors, ${T_k}^{V_{cat}}$, as categorical variables with different number of categories.

$$TNR = {T_f}^{V_{cont}} \cdot {T_k}^{V_{cat}} \tag{6}$$

For instance, if there were two continuous explanatory variables with three fuzzy sets each, and one categorical variable with five categories, TNR would be $3^2 \cdot 5^1 = 45$ rules. As shown, the number of rules that compose the rule matrix grows exponentially with the number of variables. Therefore, we decided to include the selection of variables in the GA in order to maintain the results' interpretability. In this way, the use of variables that do not influence pipe failures would be avoided.

2.1.3. Classification

The assignment of a class to each input sample x_i of the test data is done according to a rule matrix Q. Both the matching degree of the samples with the rules ($w_q(x_i)$), and the rule weights (RW_q) are employed to make the classifications. Therefore, the first step is to calculate the matching degrees of that new samples with all the rules of the rule matrix. Rule weights, which help to identify the rules that better discriminate between classes, have been previously established based on training data. The class of the rule, whose product between the matching degree and the rule weight is the maximum, is assigned to each test sample (Eq. (7)).

$$y_i = C_q \backslash \max\left\{w_q\left(x_i\right) \cdot RW_q | R_q \epsilon Q\right\} \quad \forall i \in Test\ data \tag{7}$$

2.2. Genetic algorithm

Genetic algorithms (GAs) are search algorithms inspired by Darwin's Theory of Evolution which defends that species survive through a process named *natural selection*. They were first formulated by John Holland in 1975 [39] and his disciple David Goldberg was the first to apply them to industrial problems [40].

GAs are population metaheuristics that explore the solution space in order to find the global optimum of a problem. The search process begins from a set of solutions or individuals called population. From this population, two solutions or parent chromosomes are selected and, then, crossover and mutation mechanisms are applied to generate two new solutions also named children chromosomes.

As mentioned before, there are many options to optimize MFs. This study focuses on exploring the positions of the intervals in the variables' range. For this purpose, the solution encoding is a key component that affects the proposed methodology, which can allow or avoid certain movements. In addition, we follow the criterion of maintaining the encoding as simple as possible in order to ease the interpretation. Although there are many evolutionary algorithms, GA was chosen in this study due to three main reasons: (i) its versatility to encode solutions; (ii) its exploration capabilities; and (iii) the quantity of empirical and theoretical research that use this EA to optimize fuzzy logic-based systems as it can be seen in Table 1. As a disadvantage, GAs have a lot of hyperparameters to be stated, which increases the difficulty to calibrate the system properly.

2.2.1. Individuals and population

The designed genetic algorithm aims to optimize the selection of variables and to tune membership functions. Firstly, each individual has as many binary gens as possible variables to choose. Thus, if the gen is 1, the variable participates in the fuzzy system. On the contrary, if the variable is not included in the fuzzy system, its associated gen is 0. Secondly, there is one real gen

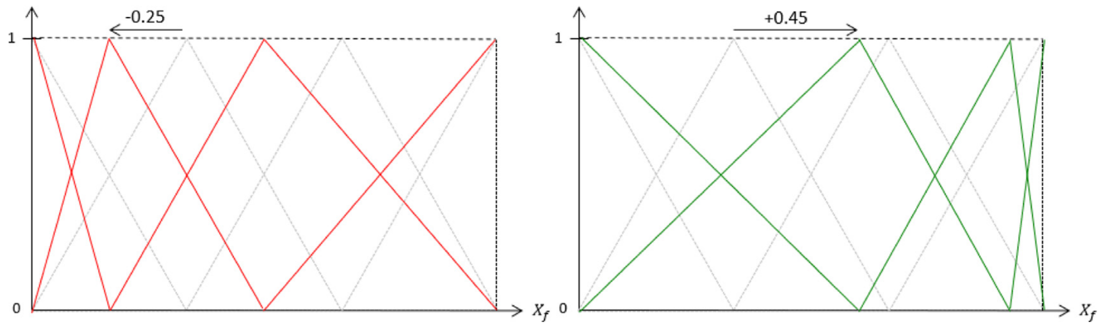

Fig. 5. Core displacement of membership functions with 4 fuzzy sets – (a) negative −0.25; (b) positive 0.45.

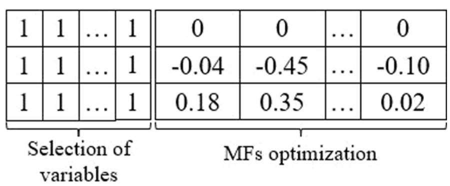

Fig. 6. Three first chromosomes of the population.

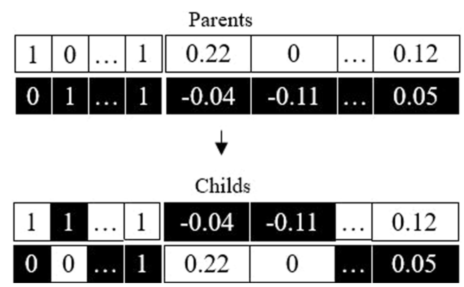

Fig. 7. Uniform crossover.

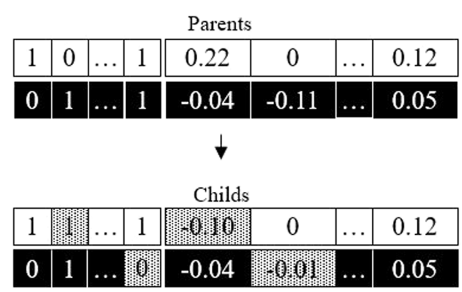

Fig. 8. Mutation process.

associated to each numerical variable which represents the core displacement of its fuzzy sets. These gens vary from −0.45 to 0.45, which means that the cores of the fuzzy sets can move to the left and to the right until 45% of the initial set width. As previously said, MFs are strong and initially uniform, so the width of fuzzy sets is directly related to the universe of discourse of each variable. This process can be better appreciated by attending to Fig. 5, which represents two core displacements, one negative and one positive, of a four-partitions' membership function. As it can be seen, negative displacements (left graph) prioritize the discrimination between low values of a variable, while positive displacements (right graph) emphasize the differences of higher values of a variable. Meanwhile, categorical variables only have binary gens that represent whether the variable is selected or not. Based on a previous work [32], three individuals are always added to the initial population (see Fig. 6), all with 1 in their binary part, which means that all variables are selected. Regarding their real parts, MFs optimization, the first individual has no core displacement, while the second one has only negative displacements and the last one, only positive displacements. The rest of the individuals that compose the population are randomly generated. The population size is set to ten or twenty chromosomes after GA calibration. The selection process is chosen between random and tournament. The tournament selection consists of selecting the two best chromosomes between four randomly chosen ones. Once crossover or mutation is applied to two parent chromosomes, two new chromosomes, also named child chromosomes, are introduced in the population at the same time as two old chromosomes are eliminated. The two eliminated chromosomes are randomly chosen, assuring not to eliminate the best chromosome of the population, which is known as elitism. In the replacement process, verifications make sure that no child is in the population yet, in order to avoid repetitions. If this happened, a new random individual is generated and included in the population.

2.2.2. Crossover and mutation

In the search for the optimum, crossover and mutation mechanisms are essential to achieve a suitable trade-off of the exploitation versus the exploration of the search space [41]. While crossover is related to exploitation, mutation concerns exploration. Although both mechanisms are important, the exploitation may be more significant to find the global optimum. Therefore, the crossover probability is often higher than the mutation probability. In this study, several crossover and mutation probabilities are tested in order to find the most suitable values.

Uniform crossover

Uniform crossover consists of interchanging gens of parent chromosomes with a probability of 0.5. Fig. 7 shows an example of the crossover operation. *Mutation*

A simple Bit Flip mutation is applied to the binary part of the chromosome, while to the real part, a Gaussian mutation is chosen (see Fig. 8). It consists of adding a random value from a Gaussian distribution. Only those real gens whose binary associated gen is 1 can mutate.

2.2.3. Fitness function

Once individuals have been designed, there is a need to define a criterion or metric to evaluate the fitness of each individual or solution. The confusion matrix, tool chosen in this study, measures the performance of classifiers by comparing the real and predicted output variable. For this purpose, test data is used, and the classification is made according to the fuzzy system generated by each solution. Several metrics derive from this matrix, including the recall or true-positive rate (Eq. (8)) and the specificity, or true-negative rate (Eq. (9)). On the one hand, TP and TN count the number of class 1 and class 0 instances that are well predicted. On the other hand, FP and FN count the number of wrong predictions of each class.

$$TP_{rate} = \frac{TP}{TP + FN} \quad (8)$$

$$TN_{rate} = \frac{TN}{TN + FP} \quad (9)$$

Even though our objective is to avoid the maximum number of pipe failures, which is to maximize TP_{rate}, companies can only replace a low percentage of pipes due to budget limitations. For this reason, the number of correct predictions of pipes which do not fail, TN_{rate} must also be maximized. As suggested by [28], the Area Under the Curve (AUC) considering a single point is an appropriate fitness function to optimize both rates (Eq. (10)). Moreover, it is not a very computationally demanding metric.

$$AUC = \frac{TP_{rate} + TN_{rate}}{2} \quad (10)$$

3. EFS implementation and results

In order to validate the usefulness and the performance of the proposed methodology, it has been implemented to a real case study. The full description of the case study and the obtained results are assumed as a significant contribution of this paper.

3.1. Case study and data processing

The data used in this study are from the water supply network of a Spanish city. This network supplies drinking water to more than 1,000,000 people from the city and the surrounding area. Data have been updated from a previous study [14]. That study employed eight explanatory variables, however, three of them showed to be far less influential than the others. These are number of connections of the pipe section, type of network and pressure fluctuation. As evolutionary fuzzy systems are high computational consuming and the number of explanatory variables increases the number of rules, these three variables have been eliminated from the present study. Therefore, the possible input variables are pipe material (MAT), diameter (DIA), age (AGE), number of previous failures (NOPF) and length (LEN).

The pipe material is a categorical variable with five different categories: ductile iron (DI), cast iron (CI), polyethylene (PE), concrete (CON) and asbestos cement (AC). Pipes of materials with a presence of less than 1% on the network have been excluded from the study. Consequently, the total network length is approximately 3700 km. In the seven years of study, 4398 pipe failures have been recorded: 223 (DI), 962 (CI), 130 (PE), 95 (CON) and 2988 (AC). Additionally, the annual failure rate per kilometer and material is presented in Fig. 9.

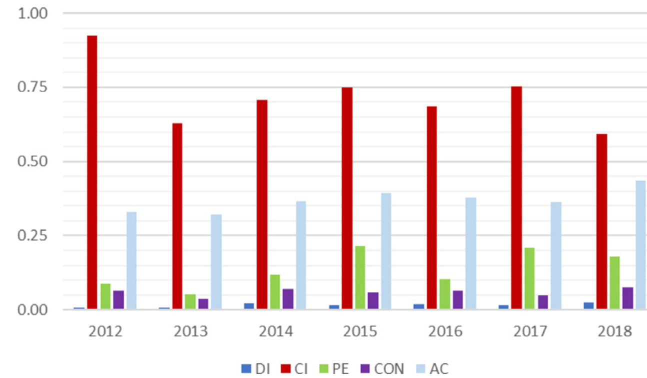

Fig. 9. Annual failure rate per kilometer and material.

Table 2
Description of numerical variables.

Var.	Units	Mean	Std	Min	Max
DIA	mm *	152.31	142.08	20	1700
AGE	years	25.91	17.16	0	119
LEN	m	42.86	79.27	0.50	4295
NOPF	–	0.05	0.31	0	11

Table 2 summarizes the main characteristics of numerical variables. Both diameter and number of previous failures are discrete variables. On the one hand, pipe diameters are established according to each material by manufacturing companies. As it can be seen in the table, the pipes diameters in the network vary from 2 to 170 centimeters, having more than 40 different diameters. On the other hand, the universe of discourse of NOPF ranges from 0 to 11, being an integer variable, i.e., there are twelve different values. In both cases, the number of possible values is high enough compared to its universe of discourse. For this reason and in order to simplify calculations, they are both treated as continuous variables.

Data need to be carefully processed to get the most out of them. In machine learning applications in the area, it is a conventional procedure to use 70% to 80% of the data to train the model and 20% to 30% to test it [42]. In fact, many studies have used this training-test configuration with other predictive systems as random forest, BBN or ANN [2,13,43,44]. In our study, the raw data are firstly extrapolated into a yearly basis. Secondly, they are divided into training set (first 5 years) and test set (last 2 years). The former is used to train the system and the latter to evaluate its performance by means of quality metrics. The test set contains 1,330 pipe failures, which corresponds to 30% of the total registered pipe failures.

The first challenge is that data from water supply networks are totally unbalanced, in our case with an order of 8:1000, which means that there are eight registered pipe failures for every thousand samples. Class imbalance causes the rule matrix to focus on the correct classification of the majority class. For this reason, many research use sampling methods, however, it inevitably introduces noise or requires a loss of valuable information. Sanz et al. [28] designed a procedure to rescale rule weights in order to avoid the need of sampling methods. In their study, the obtained results are very promising, but the size of their datasets is substantially lower than our case study. As previously stated, the training of fuzzy systems consumes much time and, as this training is iterative, the computational time increases enormously. Since our dataset is very extensive, we have decided to apply an under-sampling technique instead of the rescaling procedure proposed by Sanz et al. Nevertheless, it can be considered for future work.

3.2. Calibration of the GA

To strengthen the capabilities of the GA, it is highly recommended to calibrate its parameters. In this study, a battery of simulations is carried out to discover the best configuration for population size, crossover and mutation probabilities, and selection process. Table 3, Tables 4 and 5 present the fitness function (AUC), the TP_{rate}, the TN_{rate} and the total number of rules of each simulation. Each table contains 16 simulations for 3, 4 and 5 fuzzy sets (FSs) of numerical variables respectively. The best AUC obtained in each case is marked and its associated configuration of parameters is used to obtain the final results. The simulations have 50 generations. Furthermore, runtimes are included due to the intensive computational load of EFSs, which incites us to avoid configurations with runtimes excessively high. The code has been programmed in Python 3.7, using a system

Table 3
Battery of simulation to calibrate the GA for 3 FSs.

Sim.	1	2	3	4	5	6	7	8
Pop size	10	10	10	10	10	10	10	10
CXPB	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.8
MUTPB	0.5	0.5	0.4	0.4	0.3	0.3	0.2	0.2
Select.	Random	Tourn.	Random	Tourn.	Random	Tourn.	Random	Tourn.
Time (s)	587	401	309	312	308	390	515	659
AUC	0.76148	0.76109	0.76079	**0.76153**	0.76078	0.76081	0.76036	0.76071
TP_{rate}	0.90484	0.90739	0.90314	0.90484	0.90314	0.90314	0.90484	0.90824
TN_{rate}	0.61811	0.61479	0.61843	0.61822	0.61842	0.61847	0.61587	0.61318
TNR	45	135	45	45	45	45	405	405
Sim.	9	10	11	12	13	14	15	16
Pop size	20	20	20	20	20	20	20	20
CXPB	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.8
MUTPB	0.5	0.5	0.4	0.4	0.3	0.3	0.2	0.2
Select.	Random	Tourn.	Random	Tourn.	Random	Tourn.	Random	Tourn.
Time (s)	354	457	428	323	314	458	316	355
AUC	0.76047	0.76148	0.76053	0.76079	0.76141	0.76117	0.76042	0.76137
TP_{rate}	0.90229	0.90739	0.90229	0.90314	0.90484	0.90484	0.90229	0.90484
TN_{rate}	0.61865	0.61558	0.61877	0.61843	0.61797	0.61750	0.61855	0.61790
TNR	45	405	45	45	45	45	45	45

Table 4
Battery of simulation to calibrate the GA for 4 FSs.

Sim.	17	18	19	20	21	22	23	24
Pop size	10	10	10	10	10	10	10	10
CXPB	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.8
MUTPB	0.5	0.5	0.4	0.4	0.3	0.3	0.2	0.2
Select.	Random	Tourn.	Random	Tourn.	Random	Tourn.	Random	Tourn.
Time (s)	538	2170	2059	1693	1070	2409	1294	2585
AUC	0.76383	0.76161	0.76373	0.76233	0.76376	0.76343	0.76380	0.76395
TP_{rate}	0.91079	0.90739	0.91334	0.90909	0.91079	0.90994	0.91334	0.91419
TN_{rate}	0.61687	0.61583	0.61413	0.61557	0.61672	0.61691	0.61426	0.61371
TNR	80	1280	1280	1280	80	80	1280	1280
Sim.	25	26	27	28	29	30	31	32
Pop size	20	20	20	20	20	20	20	20
CXPB	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.8
MUTPB	0.5	0.5	0.4	0.4	0.3	0.3	0.2	0.2
Select.	Random	Tourn.	Random	Tourn.	Random	Tourn.	Random	Tourn.
Time (s)	960	936	954	1809	529	1126	975	2072
AUC	0.76154	0.76338	0.76210	0.76264	0.76384	**0.76439**	0.76376	0.76257
TP_{rate}	0.90484	0.91419	0.91249	0.91164	0.91079	0.91504	0.91079	0.90824
TN_{rate}	0.61824	0.61258	0.61170	0.61365	0.61688	0.61374	0.61672	0.61689
TNR	80	1280	80	1280	80	1280	80	320

Table 5
Battery of simulation to calibrate the GA for 5 FSs.

Sim.	33	34	35	36	37	38	39	40
Pop size	10	10	10	10	10	10	10	10
CXPB	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.8
MUTPB	0.5	0.5	0.4	0.4	0.3	0.3	0.2	0.2
Select.	Random	Tourn.	Random	Tourn.	Random	Tourn.	Random	Tourn.
Time (s)	4598	17367	10793	14113	10178	3615	1504	14559
AUC	0.76414	0.76430	0.76427	0.76413	0.76377	0.76393	0.76382	0.76384
TP_{rate}	0.91419	0.91419	0.91504	0.91419	0.91504	0.91079	0.91079	0.91419
TN_{rate}	0.61408	0.61441	0.61350	0.61407	0.61250	0.61707	0.61686	0.61350
TNR	3125	3125	3125	3125	3125	125	125	3125
Sim.	41	42	43	44	45	46	47	48
Pop size	20	20	20	20	20	20	20	20
CXPB	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.8
MUTPB	0.5	0.5	0.4	0.4	0.3	0.3	0.2	0.2
Select.	Random	Tourn.	Random	Tourn.	Random	Tourn.	Random	Tourn.
Time (s)	5463	17051	3212	8732	2198	18674	5618	5169
AUC	0.76440	0.76472	0.76392	0.76346	0.76381	**0.76510**	0.76443	0.76379
TP_{rate}	0.91504	0.91589	0.91419	0.90994	0.91079	0.91674	0.91504	0.91334
TN_{rate}	0.61377	0.61355	0.61365	0.61697	0.61683	0.61347	0.61382	0.61424
TNR	3125	3125	3125	125	125	3125	3125	3125

having Intel 7 processor and 8 GB of RAM with Windows 10 as operation system.

From the calibration of the GA, it is concluded that runtimes grow with the number of fuzzy sets, being 405 s on average for simulations with 3 partitions, 1449 s for simulations with 4 partitions, and 8928 s for simulations with 5 partitions. On the contrary, a greater population size does not imply an increase in runtimes. Regarding crossover and mutation probabilities, a clear

Table 6
Results presented as the mean and the best-attained solution of ten independent simulations – Quality metrics, selection of variables and core displacements of MFs, total number of rules (TNR) and runtimes.

FS		AUC	TP_{rate}	TN_{rate}	MAT	DIA	AGE	LEN	NOPF	TNR	Time (s)
3	Mean	0.76145	0.90467	0.61822	1	0	0	1;−0.27	1;−0.41	45–135	712
	Best sol.	0.76167	0.90569	0.61765	1	0	1; 0.02	1;−0.07	1;−0.44	135	601
4	Mean	0.76393	0.91240	0.61546	1	0	0	1;−0.20	1;−0.42	80–1280	2543
	Best sol.	0.76439	0.91504	0.61374	1	1; 0.36	1;−0.21	1;−0.27	1;−0.44	1280	1126
5	Mean	0.76442	0.91402	0.61482	1	1;−0.20	1; 0.09	1;−0.12	1;−0.31	125–3125	20426
	Best sol.	0.76524	0.91674	0.61374	1	1;−0.44	1; 0.04	1;−0.42	1;−0.19	3125	27036

tendency of improvement is noticed for higher probabilities of crossover (0.7 and 0.8) in 4 and 5 partitions simulations. Whereas in Table 3, better solutions are attained when the mutation probability grows. In all cases, tournament has demonstrated to be the most suitable selection process, as well as the use of elitism in order to maintain the best individual in the population.

Finally, the learning of the system proves to be limited. Although differences between AUCs are observed, no simulation exceeds 0.77. As a positive aspect, TP rates are excellent, they are always greater than 0.9, which means that the system prioritizes the prediction of pipe failures. Even though the accuracy of the model improves with the number of fuzzy sets, the rule matrix is smaller in simulations with only 3 FSs, and fewer rules lead to higher levels of results interpretability.

3.3. Results

Results are particularly valuable and reliable because this case study is from a large water supply network with an extensive historical pipe failure database. Table 6 presents the mean and the best-attained solution of ten independent simulations for each number of fuzzy sets. The columns include, from left to right: the aforementioned quality metrics, i.e., AUC, TP_{rate} and TN_{rate}; the selection of variables and the core displacement of their MFs; the range of the total number of rules for the ten simulations; and the runtimes on average. The individual results for each of the 30 simulations are included in Table 10 in Appendix. The maximum number of iterations is fixed at one hundred iterations per simulation and the configuration of the other GA parameters is done based on the previous calibration (marked columns of Tables 3 and 4, and 5).

The designed EFS prioritizes the correct classifications of pipe failures since TP_{rates} are much higher than TN_{rates} which can have two causes. Firstly, the fitness function gives the same importance to the percentage of well-classified samples of each class, since there are much less samples of class 1 in the test set, it is easier to increase the TP_{rate}. Secondly, the rule matrix is built based on under-sampled data, which can imply the loss of some patterns from non-failure pipes. In general, as higher the number of fuzzy sets, greater values of the fitness function (AUC) are obtained. Unlike DIA and AGE, the variables MAT, LEN and NOPF have demonstrated to be the most influential in the appearance of pipe failures since they are included in the models of all simulations.

Fig. 10 shows the evolution of the fitness function for the simulations corresponding to the three best solutions of Table 6. These graphs demonstrate the correct performance of the GA since fitness functions improve over iterations, stabilizing and converging after at the end. However, as previously mentioned, a limitation in the accuracy of the model is observed since AUCs do not exceed 0.77. In order to complete the results, Fig. 11 depicts the number of predicted pipe failures for the three best-attained solutions against the network's length to be replaced for the two testing years (2017 and 2018). These graphs are similar to the popular ROC curves [45]. The vertical axis contains the true positive rate, as can be seen in the right axis Y; however, the horizontal axis represents the cumulative network length instead of the false positive rate. This modification of the ROC curve is much more interesting for companies in charge of water supply networks than the original ones. The criterion followed to rank the pipes is the product between the rule weights and the coincidence of the samples with the rules, $w_q(x_i) \cdot RW_q$. It is the same criterion that was used to assign the class to each sample. Therefore, pipes are ranked from the highest to the lowest $w_q(x_i) \cdot RW_q$, for pipes with predicted class 1, and from the lowest to the highest $w_q(x_i) \cdot RW_q$ for pipes of class 0. It should be mentioned that the first interval of the graphics is the most important since companies only replace a low percentage of pipes per year. Although the fitness function using 5 FSs is the highest, the beginning of the rankings is better for the attained solution with 3 and 4 FSs (see Fig. 11).

3.4. Adjustment of the methodology

Since only five input variables are considered in our case study, we have decided to implement a slightly modification of the methodology. In this new approach, the selection of variables is done by brute force, which means that all possible combinations are tested independently. Furthermore, variables can have different number of fuzzy sets in a solution. This modification must decrease the runtimes because one single simulation covers three simulations of the original methodology. The architecture of the methodology only changes in the evaluation of individuals, which is shown in Fig. 12.

Eq. (11) represents the formula to calculate the number of combinations without repetitions of k elements from a group of n. In the case of five input variables, there are 31 possible combinations, it is $C_{5,5}+C_{5,4}+C_{5,3}+C_{5,2}+C_{5,1} = \binom{5}{5}+\binom{5}{4}+\binom{5}{3}+\binom{5}{2}+\binom{5}{1} = 31$. Each combination is referred as a model, so we have 31 different models in this adjustment of the methodology.

$$C_{n,k} = \binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} \tag{11}$$

As previously stated, in this approach each numerical variable can have a different number of partitions, from 3 to 5. For example, the diameter can have 3 partitions while the age has 4. To do this, the structure of individuals in the GA has been modified as can be seen in Fig. 13. The first part of the chromosome includes the number of partitions of numerical variables, and the last part the core displacement as in the original approach. The selection, crossover and mutation processes have also been adapted to this new structure.

3.4.1. Calibration of the GA for the adjusted methodology

There are many options to design the calibration of evolutionary algorithms, for instance, the adaptation of the Taguchi's parameter design which allows studying a large number of decision variables with smaller number of experiments [46], or the use of other intelligent approaches [47]. Given the nature of this methodology and the problem itself, i.e., there are many models

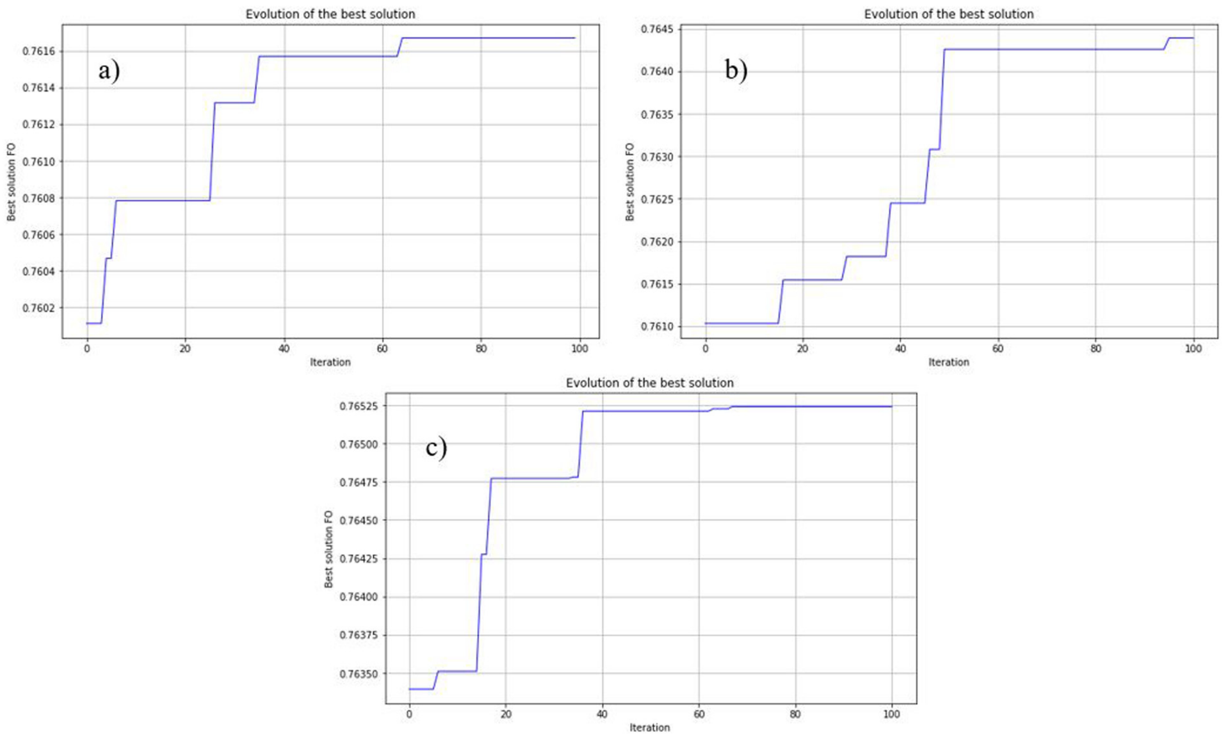

Fig. 10. Evolution of the fitness function of the three-best solution from Table 6. (a) 3 fuzzy sets; (b) 4 fuzzy sets; (c) 5 fuzzy sets.

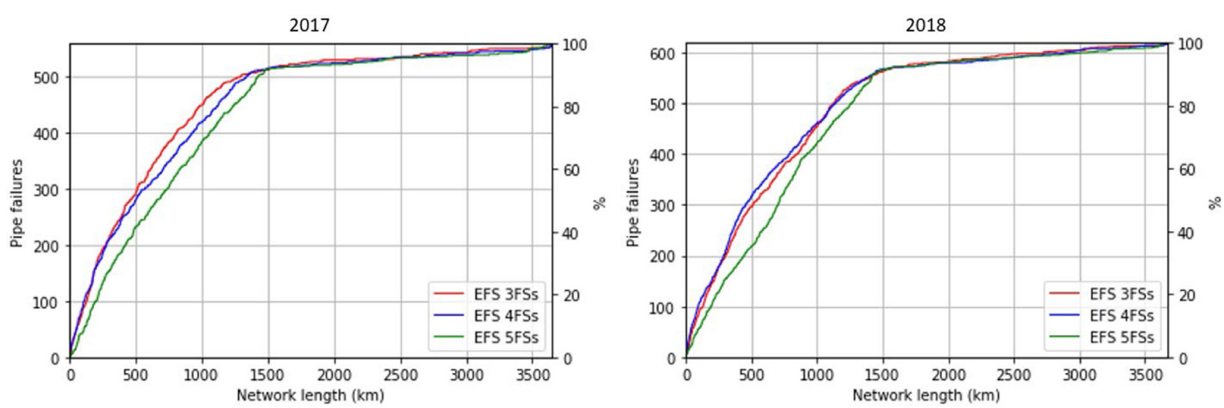

Fig. 11. Pipe failures that can be avoided according to the network length to be replaced for the three best-attained solutions of Table 6.

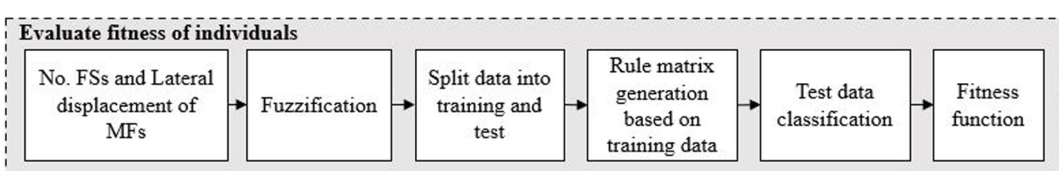

Fig. 12. Evaluation of individuals in the adjusted methodology.

5	4	…	4	0.00	0.23	…	-0.28
3	3	…	5	-0.44	0.06	…	-0.14
4	5	…	3	-0.18	0.38	…	0.03
No. FSs				MFs optimization			

Fig. 13. Three random chromosomes of the adjusted methodology.

that share multiple characteristics, in fact, they only differ in the input variables they include, in this section, we use a lab-design analysis to calibrate the GA. This strategy consists of calibrating the parameters for the most representative model, the one including all the variables, and then extending the parameters for the rest of models.

Table 7 shows the AUC, the TP_{rate}, the TN_{rate} and the total number of rules of each simulation. Following the criterion of the highest AUC, control parameters are established as: (i) pop size equals to 10; (ii) crossover and mutation probabilities of 0.8 and

Table 7
Battery of simulation to calibrate the GA for the adjusted methodology.

Sim.	1	2	3	4	5	6	7	8
Pop size	10	10	10	10	10	10	10	10
CXPB	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.8
MUTPB	0.5	0.5	0.4	0.4	0.3	0.3	0.2	0.2
Select.	Random	Tourn.	Random	Tourn.	Random	Tourn.	Random	Tourn.
Time (s)	8526	10354	8981	7580	10343	11562	11505	11106
AUC	0.76418	0.76379	0.76402	0.76436	0.76442	0.76428	0.76483	**0.76554**
TP_{rate}	0.91334	0.91334	0.91419	0.91419	0.91504	0.91419	0.91589	0.91759
TN_{rate}	0.61502	0.61425	0.61385	0.61454	0.61380	0.61437	0.61378	0.61349
TNR	1500	2500	2500	1200	2000	2500	2500	2000
Sim.	9	10	11	12	13	14	15	16
Pop size	20	20	20	20	20	20	20	20
CXPB	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.8
MUTPB	0.5	0.5	0.4	0.4	0.3	0.3	0.2	0.2
Select.	Random	Tourn.	Random	Tourn.	Random	Tourn.	Random	Tourn.
Time (s)	8664	12137	9241	8423	7511	8599	12681	8574
AUC	0.76358	0.76440	0.76399	0.76409	0.76453	0.76399	0.76395	0.76406
TP_{rate}	0.91419	0.91504	0.91419	0.91504	0.91504	0.91419	0.91419	0.91419
TN_{rate}	0.61296	0.61377	0.61379	0.61314	0.61402	0.61379	0.61372	0.61394
TNR	1500	2000	1200	960	1875	1875	2500	1200

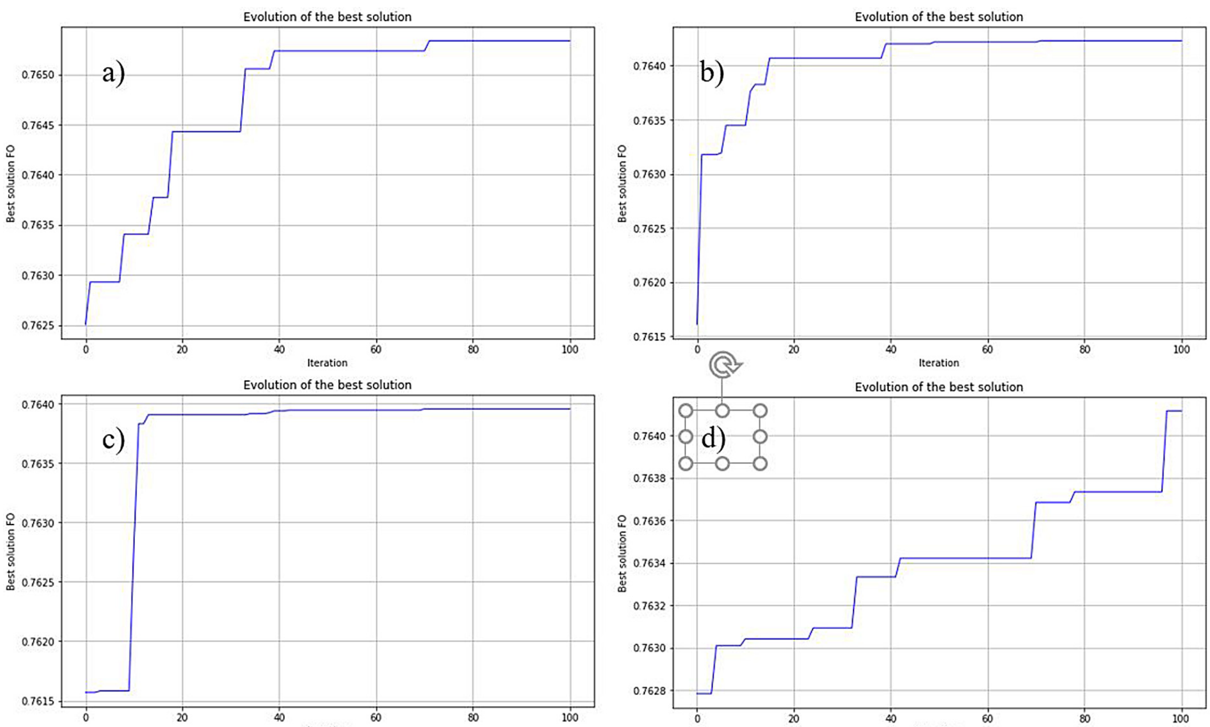

Fig. 14. Evolution of the fitness function of the four best solutions from Table 8. (a) Model 1; (b) Model 2; (c) Model 3; and (d) Model 4.

0.2 respectively; and (iii) and tournament as selection process. The simulation with this parameters' configuration (simulation 8) is given in bold in the table.

Finally, the training and testing processes are equal to the ones in the original approach, using the training data to add the consequences to the rules, and the test data to evaluate the performance of the models.

3.4.2. Results of the adjusted methodology

In order to obtain reliable results, 10 simulations for the 31 different combinations of variables (a total of 310 simulations), with 100 iterations each one, are carried out. Table 8 shows the mean and the best-obtained solution for the four models whose AUCs or fitness functions have been the maximum on average. Models are sorted according to the mean AUC of those 10 simulations i.e. model 1 is the one which has achieved the best result by average, followed by model 2, etc. The variable pipe material (MAT) is not included in the table since it is not influenced by core displacements. Results of the remaining models have been less accurate and, therefore, they have not been exposed here. Nonetheless, the best solution of each variables' combination, this achieving the highest AUC, is included in Table 11 in Appendix.

The variables included in the four models presented on Table 8 are:

- Model 1: MAT, DIA, AGE, LEN and NOPF
- Model 2: MAT, DIA, LEN and NOPF
- Model 3: MAT, LEN and NOPF
- Model 4: MAT, AGE, LEN and NOPF

Besides the quality metrics, the number of FSs of each variable and its core displacement are also presented in the table. In this methodology, the total number of rules (TNR) is directly related

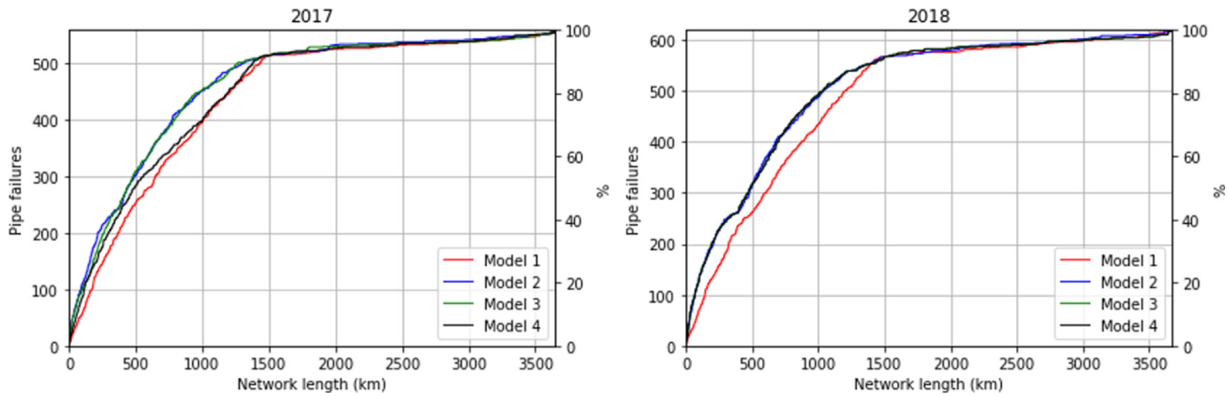

Fig. 15. Pipe failures that can be avoided according to the network length to be replaced for the four best-attained solutions of Table 8.

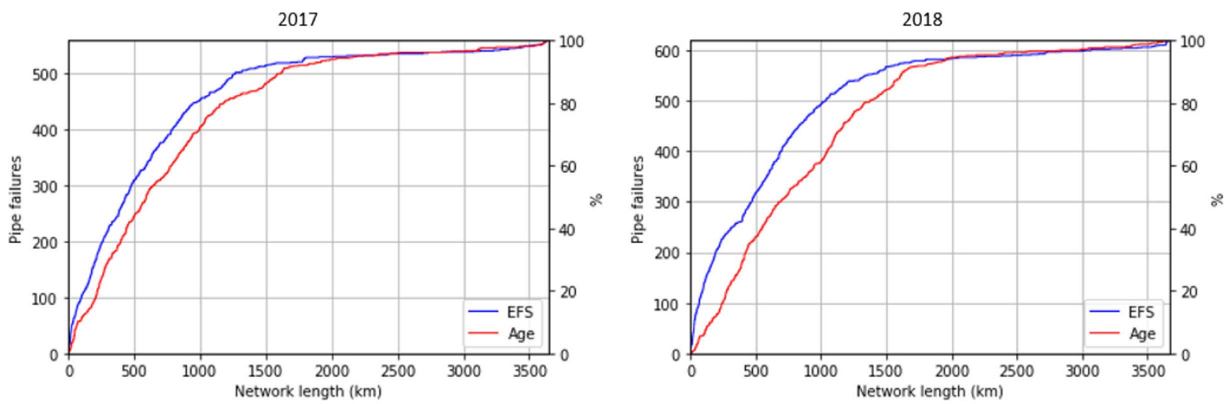

Fig. 16. Predicted pipe failures per network length – Comparison between replacement by age and by the designed EFS (Model 3 from Table 8).

Table 8
Results of the adjusted methodology presented as the mean and the best-attained solution of ten independent simulations – Quality metrics, No. FSs and core displacements of MFs, total number of rules (TNR) and runtimes.

Model	AUC	TP_{rate}	TN_{rate}	DIA	AGE	LEN	NOPF	TNR	Time (s)
Mean	0.76496	0.91573	0.61419	–	–	–	–	1125–2500	13299
Best sol.	0.76534	0.91674	0.61394	4;0.04	5;−0.05	5;−0.42	5;−0.36	2500	12660
Mean	0.76404	0.91403	0.61404	–	–	–	–	225–500	2299
Best sol.	0.76423	0.91419	0.61427	3;0.40	–	3;−0.33	5;−0.40	225	2332
Mean	0.76394	0.91079	0.61708	–	–	–	–	75–125	842
Best sol.	**0.76396**	**0.91079**	**0.61712**	**–**	**–**	**3;−0.22**	**5;−0.44**	**75**	**701**
Mean	0.76370	0.91087	0.61653	–	–	–	–	300–625	2427
Best sol.	0.76412	0.91504	0.61320	–	4;−0.36	5; 0.02	5;−0.32	500	2812

to the number of FSs. It is observed that MAT, LEN and NOPF appear in all the models, so as previously discovered with the original methodology, these variables are significant to predict pipe failures.

A slight increase of the AUC means is observed in comparison with the original methodology; therefore, this adaptation overcomes the original at least for our case study. Moreover, runtimes are substantially lower because one simulation groups three simulations of the original methodology (3, 4 and 5 FSs). It is also observed that the search for the best solution is more intensive for this configuration of the GA (see Fig. 14), since the progress of fitness functions over time shows more augmentations.

Finally, Fig. 15 shows two graphs that depict the pipes failures against the network length that should be replaced. Attending to this criterion, models 2 and 3 are the best, with model 1 showing a worse performance. This is because the fitness function only considers the percentage of right predictions but not the ranking of the pipes. For this reason, it is interesting to monitor these graphs in a post-analysis in order to select the most suitable model.

4. Discussion and qualitative analysis of rules

Comparing the original and the adjusted approaches, it can be concluded that the second one is more appropriate when there are not too many variables. According to Eq. (11), it is observed a significant increase in the number of combinations/models when the input variables increase one unit. In fact, if there are 7, 8, 9 or 10 input variables, the number of models will be 127, 255, 511 or 1023 respectively. This would substantially increase the runtimes as well as hinder the analysis of the results, making the methodology much less flexible. Consequently, the adjusted methodology is only recommended if the number of explanatory variables does not exceed 6. In other cases, the original methodology would have a better performance.

Although evolutionary algorithms cannot guarantee the obtention of the optimal solution, Figs. 10 and 14 allow us to state that the developed system achieves high quality solutions, since the fitness function improves significantly at the beginning and after that it stabilizes and converges once a number of iterations have been produced. In this section, we analyze the best-attained

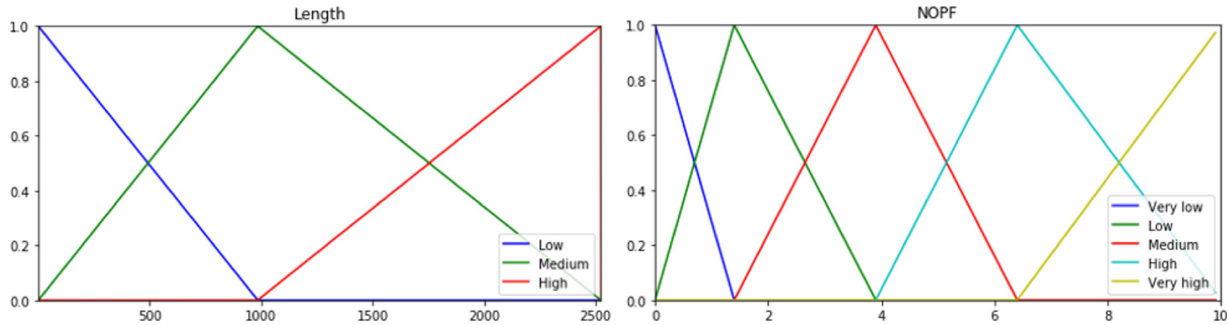

Fig. 17. Fuzzification of the numerical variables according to the best-attained solution for model 3. Core displacement of Length −0.22 and NOPF -0.44.

Table 9
The three best rules of each class according to the product Confidence*Support for the best-attained solution of Model 3 [MAT, LEN and NOPF].

Rule	$c\left(R_q \to C_q\right)$	$s\left(R_q \to C_q\right)$
IF MAT is AC and LEN is low and NOPF is very low THEN, Class is 1	0.63145	0.25620
IF MAT is CI and LEN is low and NOPF is very low THEN, Class is 1	0.75218	0.09628
IF MAT is AC and LEN is low and NOPF is low THEN, Class is 1	0.93422	0.03313
IF MAT is DI and LEN is low and NOPF is very low THEN Class is 0	0.92195	0.25279
IF MAT is PE and LEN is low and NOPF is very low THEN Class is 0	0.75272	0.03146
IF MAT is DI and LEN is medium and NOPF is very low THEN Class is 0	0.85208	0.01220

solution by the third model of Table 8. This holds a special interest because it achieves good levels of quality metrics, AUC of 0.76396 and TP_{rate} of 0.91079, with only 75 rules. Moreover, it accomplishes a successful ranking of the pipes according to $w_q(x_i) \cdot RW_q$ (Fig. 15). This solution includes three variables: MAT, LEN and NOPF. Fig. 16 depicts the pipe failures that could have been avoided using this solution and if the company had replaced pipes according to their age for years 2017 and 2018. The improvements achieved by the proposed methodology is clearly appreciated, as well as the advantages that its application can suposse to companies. For example, by replacing 250 km of pipes in 2018, which only represents 6.75% of the total network length, 228 pipe failures would be avoided (blue line). Meanwhile, following the policy of replacing pipes according to their age, only 100 pipe failures are avoided (red line).

Fig. 17 shows the fuzzification of the numerical variables according to the analyzed solution, where LEN has 3 FSs and a core displacement of −0.22, and NOPF has 5 FSs and a core displacement of −0.44.

As previously stated, the rule matrix derived from this solution has 75 rules, however, 32 of them have a support equal to 0, which means that they do not cover any training sample. From the other 43 rules, Table 9 shows the three rules of each class with the highest product between the confidence and the support of the rules with their classes, $c\left(R_q \to C_q\right) * s\left(R_q \to C_q\right)$. If we used the rule weight (RW_q) or the confidence as criterion, only rules that cover a very small number of samples would appear in the table, which is not representative. By using the aforementioned product, we show the rules that are more discriminating at the same time that cover a significant number of samples.

Since in the studied solution only three variables are selected, each rule has three antecedents. All rules whose antecedents include medium, high or very high NOPF, regardless of their material and their length, predict that the pipe is going to fail, but their supports are so small that they do not appear in the table. Therefore, it can be stated that pipes that have suffered a previous failure are more likely to fail again. This may be caused by some inattention or problem in the pipe replacement system. Therefore, it is recommended to revise the replacement processes of the company.

Regarding the material, most rules corresponding to AC and CI pipes are of Class 1, so the replacement of these materials should be prioritized. On the contrary, the non-failure rules are characterized by DI, PE and CON pipes. However, concrete pipes with low NOPF and medium LEN are likely to break. Finally, a tendency to fail in pipes with higher lengths is appreciated. The only case in which a pipe with high LEN is not predicted to fail is for DI pipes with very low NOPF.

The main advantages of the designed EFS are its ability to represent knowledge in a natural way for human understanding (rule matrix), and its learning and adaptation capabilities (evolutionary algorithm). It is very relevant to preserve the interpretability of the results since although other black box models can obtain higher levels of accuracy, the absence of a natural understanding of the methodology can lead to a lack of confidence in the process. As a disadvantage, the high specificity of rules obtained by our EFS hinders the attainment of general conclusions. Also, this technology is computationally expensive. However, this is not so important because the system supports long-term decisions regarding maintenance and replacement policies.

5. Conclusions

This study presents the design of an EFS to optimize pipe replacement plans of water supply systems. The pipe replacement policy of water companies is frequently established according to cultural or historical procedures. Consequently, some hidden causes of unexpected pipe failures go unnoticed. The proposed methodology is a perfect tool to discover these causes, so companies can gain more robust and independent decision-making systems. In fact, the great potential of fuzzy logic lies in its capacity to describe the causes and consequences of a phenomenon. The use of fuzzy logic as machine-learning system allows obtaining easily interpretable rules that characterize the problem.

In this work, a genetic algorithm is designed to establish the parameters of a fuzzy logic system based on training data. The objective function of the GA is to maximize the AUC or area under the curve considering a single point. However, more metrics as the True-Positive and True-Negative rates are also considered. It is worth mentioning that the computation of the rule matrix

Table 10
Results of the 30 simulations individually for the original methodology. Quality metrics, selection of variables {1, 0} and the core displacements of their MFs, total number of rules (TNR) and runtimes.

FS	Sim.	AUC	TP_{rate}	TN_{rate}	MAT	DIA	AGE	LEN	NOPF	TNR	Time (s)
3	1	0.76153	0.90484	0.61822	1	0	0	1;−0.30	1;−0.43	45	702
	2	0.76055	0.90229	0.61880	1	0	0	1;−0.27	1;−0.35	45	666
	3	**0.76167**	**0.90569**	**0.61765**	**1**	**0**	**1;0.02**	**1;−0.07**	**1;−0.44**	**135**	**601**
	4	0.76158	0.90484	0.61831	1	0	0	1;−0.41	1;−0.43	45	625
	5	0.76155	0.90484	0.61825	1	0	0	1;−0.33	1;−0.42	45	634
	6	0.76149	0.90484	0.61814	1	0	0	1;−0.17	1;−0.41	45	639
	7	0.76156	0.90484	0.61827	1	0	0	1;−0.35	1;−0.42	45	750
	8	0.76150	0.90484	0.61815	1	0	0	1;−0.20	1;−0.40	45	625
	9	0.76157	0.90484	0.61830	1	0	0	1;−0.40	1;−0.42	45	1175
	10	0.76147	0.90484	0.61810	1	0	0	1;−0.15	1;−0.40	45	701
4	11	0.76385	0.91079	0.61692	1	0	0	1;−0.05	1;−0.45	80	1010
	12	0.76384	0.91079	0.61688	1	0	0	1;−0.08	1;−0.41	80	3239
	13	0.76383	0.91334	0.61431	1	1; 0.00	1;−0.05	1;−0.42	1;−0.40	1280	4237
	14	0.76389	0.91334	0.61444	1	1;−0.17	1;−0.40	1;−0.39	1;−0.38	1280	1849
	15	0.76385	0.91079	0.61691	1	0	0	1;−0.12	1;−0.42	80	3534
	16	0.76385	0.91079	0.61692	1	0	0	1;−0.05	1;−0.45	80	1670
	17	0.76404	0.91419	0.61389	1	1; 0.01	1;−0.27	1;−0.24	1;−0.43	1280	1970
	18	0.76383	0.91079	0.61687	1	0	0	1; 0.02	1;−0.38	80	1882
	19	0.76394	0.91419	0.61368	1	1; 0.22	1;−0.21	1;−0.37	1;−0.41	1280	3369
	20	**0.76439**	**0.91504**	**0.61374**	**1**	**1;0.36**	**1;−0.21**	**1;−0.27**	**1;−0.44**	**1280**	**1126**
5	21	0.76391	0.91079	0.61704	1	0	0	1; 0.32	1;−0.43	125	14697
	22	**0.76524**	**0.91674**	**0.61374**	**1**	**1;−0.44**	**1; 0.04**	**1;−0.42**	**1;−0.19**	**3125**	**27036**
	23	0.76437	0.91504	0.61371	1	1;−0.2	1; 0.02	1;−0.31	1;−0.40	3125	27564
	24	0.76521	0.91674	0.61369	1	1;−0.13	1; 0.01	1;−0.40	1;−0.16	3125	25659
	25	0.76391	0.91079	0.61704	1	0	0	1; 0.43	1;−0.41	125	7781
	26	0.76413	0.91419	0.61408	1	1;−0.19	1; 0.17	1;−0.12	1;−0.18	3125	21475
	27	0.76482	0.91589	0.61376	1	1;−0.32	1; 0.25	1;−0.41	1;−0.29	3125	35916
	28	0.76462	0.91504	0.61420	1	1;−0.03	1;−0.07	1;−0.38	1;−0.17	3125	28211
	29	0.76391	0.91079	0.61703	1	0	0	1; 0.25	1;−0.44	125	5589
	30	0.76406	0.91419	0.61394	1	1;−0.07	1;0.22	1;−0.19	1;−0.40	3125	10334

Table 11
Best solution obtained for each model of the adjusted methodology. Quality metrics, No. FSs and core displacements of MFs for each variable, total number of rules (TNR) and runtimes.

Model	AUC	TP_{rate}	TN_{rate}	MAT	DIA	AGE	LEN	NOPF	TNR	Time (s)
1	0.76534	0.91674	0.61394		4; 0.04	5;−0.05	5;−0.42	5;−0.36	2500	12660
2	0.76335	0.87511	0.65559	–	5; 0.03	4;−0.37	5;−0.43	5; 0.42	500	1866
3	0.76412	0.91504	0.61320		–	4;−0.36	5; 0.02	5;−0.32	500	2812
4	0.75936	0.82328	0.69544	–	–	5; 0.04	5;−0.40	4;−0.31	100	829
5	0.76326	0.91419	0.61233		3;−0.33	5; 0.01	–	5;−0.12	375	2184
6	0.74858	0.82923	0.66793	–	5;−0.41	5;−0.33	–	5;−0.20	125	844
7	0.76238	0.91334	0.61141		–	4;−0.41	–	5;−0.28	100	901
8	0.74491	0.83178	0.65804	–	–	4;−0.25	–	5;−0.09	20	316
9	0.75857	0.90144	0.61570		4; 0.17	5; 0.07	5;−0.43	–	500	2414
10	0.75336	0.87766	0.62907	–	4;−0.43	4;−0.41	5;−0.42	–	80	592
11	0.75822	0.90059	0.61584		–	4;−0.20	4;−0.44	–	80	721
12	0.74742	0.86151	0.63332	–	–	4;−0.39	5;−0.45	–	20	446
13	0.75511	0.89635	0.61388		3;−0.32	5; 0.00	–	–	75	706
14	0.73638	0.89635	0.57642	–	4;−0.45	4;−0.43	–	–	16	337
15	0.75401	0.89465	0.61336		–	5;−0.02	–	–	25	383
16	0.72805	0.87766	0.57845	–	–	4;−0.40	–	–	4	250
17	0.76423	0.91419	0.61427		3; 0.40	–	3;−0.33	5;−0.40	225	2332
18	0.70463	0.61937	0.78988	–	4;−0.29	–	5;−0.32	3; 0.40	60	581
19	0.76396	0.91079	0.61712		–	–	3;−0.22	5;−0.44	75	701
20	0.67983	0.53016	0.82950	–	–	–	5;−0.45	3;0.44	15	452
21	0.76235	0.91419	0.61050		3;−0.28	–	–	5;−0.20	75	658
22	0.65370	0.38403	0.92337	–	5;−0.18	–	–	3; 0.29	15	320
23	0.76156	0.91759	0.60554		–	–	–	5, −0.14	25	404
24	0.63867	0.30161	0.97572	–	–	–	–	5, 0.45	5	258
25	0.75754	0.89720	0.61789		3; 0.39	–	4;−0.43	–	60	713
26	0.68423	0.64826	0.72021	–	3; 0.20	–	3; 0.18	–	9	307
27	0.75676	0.89465	0.61886		–	–	5;−0.28	–	25	400
28	0.62845	0.44775	0.80916	–	–	–	5;−0.44	–	5	254
29	0.75424	0.89550	0.61297		5; 0.31	–	–	–	25	384
30	0.57808	0.72557	0.43058	–	5; 0.14	–	–	–	5	252
31	0.75381	0.89975	0.60787		–	–	–	–	5	15

generated by each individual in the evolutionary process requires the scan of the overall training set. Consequently, computational times are extremely high. Seeking to reduce these runtimes and taking into account that the number of input variables is not too high, the methodology is slightly modified. This adjustment helps to enhance the accuracy of the results.

Data from a Spanish city are used to demonstrate the performance of the system. Firstly, this is calibrated by a large number of simulations. Secondly, a new battery of simulations is run to extract reliable results. The rule matrix of the most successful solution has 75 rules and it achieves an AUC of 0.764 with only three explanatory variables: the pipe material, the pipe length and the number of previous failures. Representing the scope of the methodology, the study shows that by replacing 6.75% of the network length, it would have been possible to prevent 288 pipe failures in 2018, which accounts for 41.14% of the total. Finally, the analysis of rules shows the necessity of revising the pipe replacement system, since most pipes that have suffered a failure are predicted to fail again. Additionally, AC and CI materials demonstrate a really bad performance, so the priority must be to replace these pipes first.

Some future lines of research are proposed below: (i) the use of multi-objective instead of single objective optimization, penalizing the number of rules that compose the rule matrix; and (ii) the introduction of a rescaling method for the rule weights in order to avoid under-sampling as suggested by [28]. Moreover, the inclusion of new variables is also contemplated for future research, as well as the use of a different evolutionary algorithm, e.g. PSO which, unlike GA, has few parameters to adjust.

CRediT authorship contribution statement

Alicia Robles-Velasco: Investigation, Data curation, Formal analysis, Software, Writing – original draft. **Jesús Muñuzuri:** Supervision, Validation. **Luis Onieva:** Resources, Project administration. **Pablo Cortés:** Conceptualization, Writing – reviewing and editing, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors wish to acknowledge EMASESA, *Empresa Metropolitana de Abastecimiento y Saneamiento de Aguas de Sevilla*, and the Universidad de Sevilla (VI PPIT-US) for their financial support through the Distinguished Chair in Water Network Management (Cátedra del Agua EMASESA-US).

Appendix

Table 10 includes the results of the 30 simulations from which comes the data of Table 6. In the original methodology, the number of fuzzy sets (FS) varies from 3 to 5, so three independent batteries of simulations are carried out. The rows containing the best solutions of each battery are in bold.

In the adjustment of the methodology, there are 31 different models, one for each combination of variables. In order to obtain reliable results, 10 simulations are carried out for each model i.e. a total of 310 simulations. Given the extension on these results and for the sake of conciseness, only the simulations with the highest AUC for each model are presented in Table 11. The non-included variables of each model are marked with indents '–'. It should be noted that, as pipe material is a categorical variable, the number of FSs does not vary nor do core displacements occur. Finally, the rows containing the models presented in Table 8 are in bold.

References

[1] United Nations Development Programme, Human Development Report 2019: Beyond income, beyond averages, beyond today, 2019.
[2] G. Kabir, S. Tesfamariam, R. Sadiq, Predicting water main failures using Bayesian model averaging and survival modelling approach, Reliab. Eng. Syst. Saf. 142 (2015) 498–514, http://dx.doi.org/10.1016/j.ress.2015.06.011.
[3] P. Lin, X.X. Yuan, A two-time-scale point process model of water main breaks for infrastructure asset management, Water Res. (2019) 296–309, http://dx.doi.org/10.1016/j.watres.2018.11.066.
[4] A. Debón, A. Carrión, E. Cabrera, H. Solano, Comparing risk of failure models in water supply networks using ROC curves, Reliab. Eng. Syst. Saf. 95 (1) (2010) 43–48, http://dx.doi.org/10.1016/j.ress.2009.07.004.
[5] A. Robles-Velasco, P. Cortés, J. Muñuzuri, E. Barbadilla-Martín, Aplicación de la regresión logística para la predicción de roturas de tuberías en redes de abastecimiento de agua, Dir. Y Organ. 70 (2020) 78–85, http://dx.doi.org/10.37610/dyo.v0i70.570.
[6] J. Yan, et al., Towards effective prioritizing water pipe replacement and rehabilitation, in: IJCAI Int. Jt. Conf. Artif. Intell., 2013, pp. 2931–2937.
[7] A. Robles-Velasco, P. Cortés, J. Muñuzuri, L. Onieva, Estimation of a logistic regression model by a genetic algorithm to predict pipe failures in sewer networks, OR Spectrum (2021) 0123456789, http://dx.doi.org/10.1007/s00291-020-00614-9.
[8] A.F. Royce, D.G. Seth, L. Henneman, Bayesian belief networks for predicting drinking water distribution system pipe breaks, Reliab. Eng. Syst. Saf. 130 (2014) 1–11, http://dx.doi.org/10.1016/j.ress.2014.04.024.
[9] K. Tang, D.J. Parsons, S. Jude, Comparison of automatic and guided learning for Bayesian networks to analyse pipe failures in the water distribution system, Reliab. Eng. Syst. Saf. 186 (2019) 24–36, http://dx.doi.org/10.1016/j.ress.2019.02.001.
[10] M. Kutyłowska, Forecasting failure rate of water pipes, Water Sci. Technol. Water Supply 19 (1) (2018) 264–273, http://dx.doi.org/10.2166/ws.2018.078.
[11] A. Shirzad, M. Tabesh, R. Farmani, A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks, KSCE J. Civ. Eng. 18 (4) (2014) 941–948, http://dx.doi.org/10.1007/s12205-014-0537-8.
[12] R. Jafar, I. Shahrour, I. Juran, Application of Artificial Neural Networks (ANN) to model the failure of urban water mains, Math. Comput. Model. 51 (2010) 1170–1180, http://dx.doi.org/10.1016/j.mcm.2009.12.033.
[13] A.M.A. Sattar, Ö.F. Ertuğrul, B. Gharabaghi, E.A. McBean, J. Cao, Extreme learning machine model for water network management, Neural Comput. Appl. 31 (1) (2019) 157–169, http://dx.doi.org/10.1007/s00521-017-2987-7.
[14] A. Robles-Velasco, P. Cortés, J. Muñuzuri, L. Onieva, Prediction of pipe failures in water supply networks using logistic regression and support vector classification, Reliab. Eng. Syst. Saf. 196 (2020) http://dx.doi.org/10.1016/j.ress.2019.106754.
[15] J. Cózar, A. Fernández, F. Herrera, J.A. Gámez, A metahierarchical rule decision system to design robust fuzzy classifiers based on data complexity, IEEE Trans. Fuzzy Syst. 27 (4) (2019) 701–715, http://dx.doi.org/10.1109/TFUZZ.2018.2866967.
[16] M. Al-Zahrani, A. Abo-Monasar, R. Sadiq, Risk-based prioritization of water main failure using fuzzy synthetic evaluation technique, J. Water Supply Res. Technol. - AQUA 65 (2) (2016) 145–161, http://dx.doi.org/10.2166/aqua.2015.051.
[17] N.M. Amaitik, C.D. Buckingham, Developing a hierarchical fuzzy rule-based model with weighted linguistic rules: A case study of water pipes condition prediction, in: Proc. Comput. Conf. 2017, 2018, pp. 30–40, http://dx.doi.org/10.1109/SAI.2017.8252078.
[18] S. Salehi, M. Jalili Ghazizadeh, M. Tabesh, A comprehensive criteria-based multi-attribute decision-making model for rehabilitation of water distribution systems, Struct. Infrastruct. Eng. 14 (6) (2018) 743–765, http://dx.doi.org/10.1080/15732479.2017.1359633.
[19] H. Fares, T. Zayed, Risk assessment for water mains using fuzzy approach, in: Construction Research Congress, 2009, pp. 1125–1134.
[20] M.S. Islam, R. Sadiq, M.J. Rodriguez, H. Najjaran, A. Francisque, M. Hoorfar, Evaluating water quality failure potential in water distribution systems: A fuzzy-TOPSIS-OWA-based methodology, Water Resour. Manag. 27 (7) (2013) 2195–2216, http://dx.doi.org/10.1007/s11269-013-0283-6.
[21] Y. Kleiner, R. Sadiq, B. Rajani, Modeling failure risk in buried pipes using fuzzy Markov deterioration process, Pipeline Eng. Constr. (2004) 1–12, http://dx.doi.org/10.1061/40745(146)7.
[22] B. Dennis, S. Muthukrishnan, AGFS: Adaptive genetic fuzzy system for medical data classification, Appl. Soft Comput. J. 25 (2014) 242–252, http://dx.doi.org/10.1016/j.asoc.2014.09.032.
[23] E. Van Broekhoven, V. Adriaenssens, B. De Baets, Interpretability-preserving genetic optimization of linguistic terms in fuzzy models for fuzzy ordered classification: An ecological case study, Internat. J. Approx. Reason. 44 (1) (2007) 65–90, http://dx.doi.org/10.1016/j.ijar.2006.03.003.

[24] R. Alcalá, Y. Nojima, F. Herrera, H. Ishibuchi, Multiobjective genetic fuzzy rule selection of single granularity-based fuzzy classification rules and its interaction with the lateral tuning of membership functions, Soft Comput. 15 (12) (2011) 2303–2318, http://dx.doi.org/10.1007/s00500-010-0671-2.
[25] P. Ganesh Kumar, T. Aruldoss Albert Victoire, P. Renukadevi, D. Devaraj, Design of fuzzy expert system for microarray data classification using a novel genetic swarm algorithm, Expert Syst. Appl. 39 (2) (2012) 1811–1821, http://dx.doi.org/10.1016/j.eswa.2011.08.069.
[26] A. Fernández, V. López, M.J. Del Jesus, F. Herrera, Revisiting Evolutionary Fuzzy Systems: Taxonomy, applications, new trends and challenges, Knowl.-Based Syst. 80 (2015) 109–121, http://dx.doi.org/10.1016/j.knosys.2015.01.013.
[27] A. Ferranti, F. Marcelloni, A. Segatori, M. Antonelli, P. Ducange, A distributed approach to multi-objective evolutionary generation of fuzzy rule-based classifiers from big data, Inf. Sci. (Ny) 415–416 (2017) 319–340, http://dx.doi.org/10.1016/j.ins.2017.06.039.
[28] J.A. Sanz, D. Bernardo, F. Herrera, H. Bustince, H. Hagras, A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data, IEEE Trans. Fuzzy Syst. 23 (4) (2015) 973–990, http://dx.doi.org/10.1109/TFUZZ.2014.2336263.
[29] X. Xie, D. Yue, C. Peng, Relaxed real-time scheduling stabilization of discrete-time Takagi – Sugeno fuzzy systems via an alterable-weights-based ranking switching mechanism, IEEE Trans. Fuzzy Syst. 26 (6) (2018) 3808–3819, http://dx.doi.org/10.1109/TFUZZ.2018.2849701.
[30] D. Meng, Z. Pei, Extracting linguistic rules from data sets using fuzzy logic and genetic algorithms, Neurocomputing 78 (1) (2012) 48–54, http://dx.doi.org/10.1016/j.neucom.2011.05.029.
[31] M. Antonelli, P. Ducange, F. Marcelloni, A fast and efficient multi-objective evolutionary learning scheme for fuzzy rule-based classifiers, Inf. Sci. (Ny) 283 (2014) 36–54, http://dx.doi.org/10.1016/j.ins.2014.06.014.
[32] F. Aghaeipoor, M.M. Javidi, MOKBL+MOMs: An interpretable multi-objective evolutionary fuzzy system for learning high-dimensional regression data, Inf. Sci. (Ny) 496 (2019) 1–24, http://dx.doi.org/10.1016/j.ins.2019.04.035.
[33] N.L. Tsakiridis, J.B. Theocharis, P. Panagos, G.C. Zalidis, An evolutionary fuzzy rule-based system applied to the prediction of soil organic carbon from soil spectral libraries, Appl. Soft Comput. J. 81 (2019) http://dx.doi.org/10.1016/j.asoc.2019.105504.
[34] O. Kisi, Z.M. Yaseen, The potential of hybrid evolutionary fuzzy intelligence model for suspended sediment concentration prediction, Catena 174 (2019) 11–23, http://dx.doi.org/10.1016/j.catena.2018.10.047.
[35] M.E. Turan, M.A. Yurdusev, Predicting monthly river flows by genetic fuzzy systems, Water Resour Manag. 28 (13) (2014) 4685–4697, http://dx.doi.org/10.1007/s11269-014-0767-z.
[36] L. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[37] E.H. Mamdani, S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, Int. J. Man. Mach. Stud. 7 (1) (1975) 1–13, http://dx.doi.org/10.1109/TSMC.1985.6313399.
[38] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, IEEE Trans. Syst. Man Cybern. 15 (1) (1985) 116–132, http://dx.doi.org/10.1109/TSMC.1985.6313399.
[39] J.H. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, 1975.
[40] D.E. Goldberg, Genetic Algorithms in Search Optimization and Machine Learning, 1989.
[41] F. Herrera, M. Lozano, Fuzzy adaptive genetic algorithms: Design, taxonomy, and future directions, Soft Comput. 7 (8) (2003) 545–562, http://dx.doi.org/10.1007/s00500-002-0238-y.
[42] A. Robles-velasco, J. Muñuzuri, L. Onieva, M. Rodrí guez palero, Trends and applications of machine learning in water supply networks management, J. Ind. Eng. Manag. 14 (1) (2021) 45–54, http://dx.doi.org/10.3926/jiem.3280.
[43] R. Tavakoli, A. Sharifara, M. Najafi, Prediction of pipe failures in wastewater networks using random forest classification, Pipelines 2019 (2020) 90–102, http://dx.doi.org/10.1061/9780784483206.011.
[44] Z. Almheiri, M. Meguid, T. Zayed, Intelligent approaches for predicting failure of water mains, J. Pipeline Syst. Eng. Pract. 11 (4) (2020) 1–15, http://dx.doi.org/10.1061/(ASCE)PS.1949-1204.0000485.
[45] T. Fawcett, An introduction to ROC analysis, Pattern Recognit. Lett. 27 (8) (2006) 861–874, http://dx.doi.org/10.1016/j.patrec.2005.10.010.
[46] J.T. Tsai, T.K. Liu, J.H. Chou, Hybrid Taguchi-genetic algorithm for global numerical optimization, IEEE Trans. Evol. Comput. 8 (4) (2004) 365–377, http://dx.doi.org/10.1109/TEVC.2004.826895.
[47] Y. Dai, P. Zhao, A hybrid load forecasting model based on support vector machine with intelligent methods for feature selection and parameter optimization, Appl. Energy 279 (September) (2020) 115332, http://dx.doi.org/10.1016/j.apenergy.2020.115332.