Available online at www.sciencedirect.com
ScienceDirect
Energy Reports 7 (2021) 184–192
www.elsevier.com/locate/egyr

17th International Symposium on District Heating and Cooling, Nottingham Trent University, 2021, 6–9 September 2021, Nottingham, United Kingdom

District heating network maintenance planning optimization

Matteo Pozzi[a,*], Andrea Bettinelli[a], Fabrizio Detassis[a], Ettore Filippini[b], Simone Graziani[a], Stefano Morgione[a], Daniele Vigo[c]

[a] *Optit Srl, via Mazzini, 82 - 40138 Bologna (BO), Italy*
[b] *A2A, Calore e Servizi Srl, Via Lamarmora 230, 25124 Brescia (BS), Italy*
[c] *DEI & CIRI ICT, Alma Mater Studiorum Università di Bologna, viale Risorgimento, 2 - 40136 Bologna, Italy*

Received 6 August 2021; accepted 22 August 2021

Abstract

To ensure the correct functioning of district heating networks and minimize critical failures, utilities allocate every year a significant part of their budget to maintenance operations. In the present work we describe a risk-based approach implemented to tackle the problem of designing optimal multi-year maintenance campaigns, applied to the Italian city of Brescia, showing how data-driven techniques can help decision makers assess the long terms impacts of budget allocations.

Peer-review under responsibility of the scientific committee of the 17th International Symposium on District Heating and Cooling, 2021.

Keywords: District heating network; Maintenance planning; Optimization; Machine learning

1. Introduction

District heating (DH) networks represent critical assets, both for the economic investment involved in their construction and for the impact on served consumers in case of failures. Proper maintenance to keep such assets in good operative conditions is of paramount importance, yet approaches to handle long-term maintenance planning represent a significant challenge since, without being able to inspect the assets, it is hard to decide the most adequate level of intervention to keep the overall system up to standards. DH utilities' managers allocate every year a budget for maintenance activities that include planned interventions, to substitute old or critical sections of the network, and emergency interventions to resolve leaks and other incidents. From a strategic viewpoint, scarce ordinary maintenance would lead, over time, in a surge of emergency calls, potentially undermining the short-terms savings on invested capital. Reversely, one may opt for large retrofitting investments to minimize emergency interventions. This study is aimed at finding a balance between these conflicting decision drivers.

The topic of maintenance management and planning has been studied for decades. Of particular interest is the role that the various types of strategies, namely *reactive*, *preventive* and *predictive maintenance,* play in the industry's

* Corresponding author.
E-mail address: matteo.pozzi@optit.net (M. Pozzi).

https://doi.org/10.1016/j.egyr.2021.08.156

Peer-review under responsibility of the scientific committee of the 17th International Symposium on District Heating and Cooling, 2021.

landscape [1,2]. In this context, Sernhed et al. [3] shows how risk management, specifically in district heating networks, can aid in the planning of reinvestment needs. To this end they defined a set of parameters (age of pipes, customer sensitivity, etc.), that should be taken into account and used to construct a risk matrix to characterize the impact and probability of failure events. Similarly, Khan et al. [4] presents a general framework for maintenance planning. This approach makes use of various factors such as financial and human health losses along with the probability of a failure's occurrence to schedule maintenance operations accordingly. Other notable examples of risk-based maintenance planning can be found in [5–7], and [8]. These studies focused respectively on subsea pipelines, urban water distribution networks, wind turbine farms, and railways.

The present study describes the approach implemented to tackle the problem of designing 5-year maintenance campaigns for the DH network of Brescia, Italy. The solution relies on a statistical approach for risk analysis, which in turn is used to feed an optimization model, that defines the optimal maintenance plan.

This work is structured as follows: Section 2 describes the case study of the district heating network of Brescia and the data used in the analysis; Section 3 details the approach implemented to optimize the maintenance plans; Section 4 presents the main results of the study and, Section 5 draws the conclusions of this work.

2. The case study

A2A Calore & Servizi (ACS) is one of the leading Italian DH utilities, managing (amongst others in Lombardy) the Brescia DH system. The network was first laid down in 1972, which means that systematic retrofitting and substitution of old pipes (and associated infrastructures) with new ones represents both a critical challenge and a standard practice. Good management practices made so, that every single intervention since the late 70s was recorded in details and, more recently, digitalized, paving the way to the current study.

2.1. Overview of the network

The network features over 650 km long pipelines, reaching more than 20,000 consumer points and covering around 60% of the heat demand of the city. Given the system's age, a wide range of piping technologies can be found, from newer pre-insulated pipes (*PR*) to older traditional hooded pipes, characterized by different installation types, respectively referred to as *TR(traditional, non-pre-insulated steel)*, *GF(fiber–cement sheath)*, *SCT(Copper Coating Stainless Steel*) and *W (asbestos/wanit)*. Fig. 1 (left) shows the current composition of the network as a breakdown of the number of pipes by technology and installation year. For the most part, the current network is composed by pre-insulated pipes, which was the emerging technology in the '80s, when the largest part of the network was built. However, other types of technologies still coexist, specifically TR, GF, W and, even if in very small numbers, SCT. Another element to be taken into account was the distribution of operational manholes, which require expensive upkeep operations.

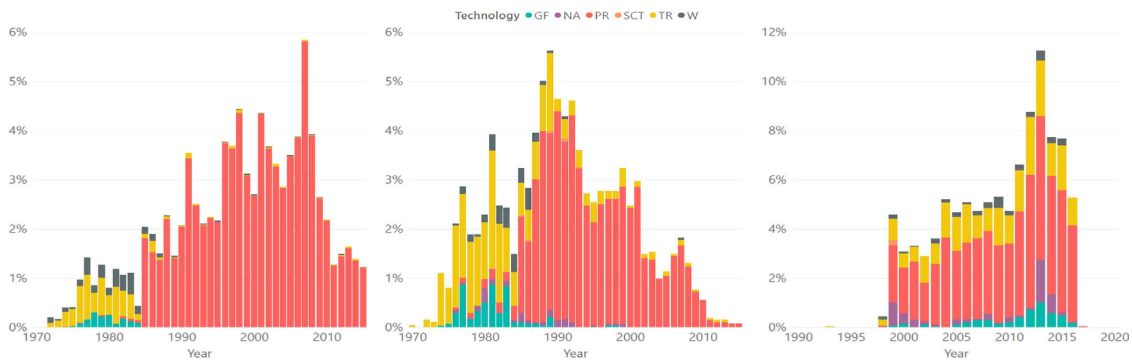

Fig. 1. Composition of the network as number pipes by technology and installation year (left); percentage of failure events by technology and installation year (middle); percentage of failure events by technology and failure's year (right).

2.2. Network's failures

While the accurate knowledge of the composition of the network is obviously valuable, it cannot provide, alone, definitive information about potential future failures, which is key for long-term planning of maintenance operations. To this end, we used the available records of failure events from the year 1999 to the year 2016, that ACS had been recording. Each failure event is characterized by the year of occurrence, the year of the initial installation of the affected pipe, its geographical coordinates, the technology, and other physical properties. This data represents only a partial representation of the full evolution of the network, since pipes could either be fixed or replaced and most records specify the occurrence of the event, but not necessarily how it was addressed. Yet, for simplicity's sake, it can be assumed that each failure originated a replacement that fully restored functionalities, as if the asset was fully renewed.

Fig. 1 (middle and right) shows a breakdown of the recorded failure events by technology and either installation year (middle) or occurrence year (right). It is noticeable how, even though most failure events are related to the pre-insulated pipes (as it can be expected, given the technology mix), every year a number of failures occur on older, non-PR pipes. The evolution of the network can also be seen by the geographic distribution of failures (Fig. 2), that adopts a similar color code to link the interventions to the piping technology.

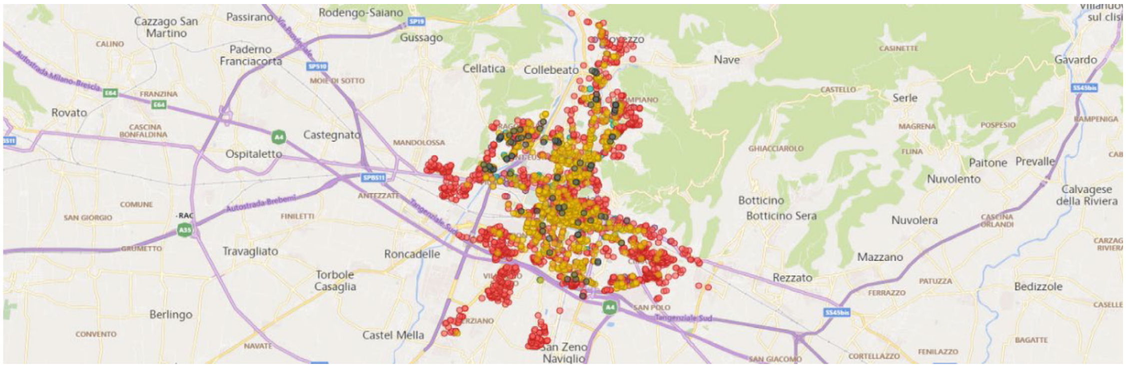

Fig. 2. Geographic distribution of the recorded failure events.

3. Approach

Given a cartographic representation of the current network and a dataset of recorded failures, the problem of maintenance planning has been tackled with an approach made up by three phases:

(i) The n*etwork's aggregation* phase takes in input the current network's structure and creates a simpler and more homogeneous representation, mostly related to data management.
(ii) Using the aggregated network and the recorded failures, the *risk analysis* phase defines the means to quantify the risk of failure of the single pipes.
(iii) Finally, the *maintenance optimization* phase determines the best long-term plan that minimizes the costs and satisfies the given constraints.

In the following, each one of these processes is described in detail.

3.1. Network's aggregation

From a cartographic data point of view, the network described provided by ACS was too fine-grained and heterogenous, both for analysis and maintenance planning purposes. Thus, the network was homogenized by coalescing adjacent pipes with compatible characteristics (i.e., same technology, diameter, etc.) and subsuming customers' feeding branches by their respective backbone pipes, bringing the original network from ∼30,000 pipes to ∼5000 pipes, allowing a much more manageable data set.

3.2. Risk analysis

The purpose of the risk analysis phase is to provide the means to quantify the risk of failures of the pipes in the network. Indeed, to be able to properly plan maintenance operations, the impact of the failures must be coupled with their likelihood, so that operations can be scheduled only when needed.

To achieve this goal, we used the failures data described in Section 2.2. In particular, the features considered for the analysis are: the installation year, the technology, the diameter, and the geographical risk of the pipe. The geographical risk is the result of a separate analysis performed to assign a score to different areas of the city of Brescia. Notice that this dataset presents some characteristics that might hinder the analysis process: while we can observe the current state of the network, we have records of the failures only from year 1999 to year 2016. This means that:

- There can be both cases of pipes that have been observed from the installation time to the failure time, and cases where installed pipes are still up and running; the latter can be referred as *right-censored* [9].
- Similarly, we might have cases of pipes where both installation year and failure year fall outside the observed period of the dataset; these cases are referred as *left-truncated* [9].
- Finally, the distribution of both technology and diameter variables *might* have changed over time, for example, early installations were made using different technologies.

Given this dataset, a variety of approaches can be applied to risk analysis, mainly based on machine learning techniques or survival analysis (stochastic survival models) [9]. We opted for the latter, which are specifically made to deal with the kind of data at our disposal, are easier to interpret, and have already been used in a similar case (see [10]). A discussion about when certain approaches are better than others is out of scope of this work, but we refer to Wang et al. [11] for more details.

The basic equations of survival analysis are described below: given the probability density function (p.d.f.) $f(\cdot)$ of failure events, the cumulative distribution function (c.d.f.) $F(\cdot)$ is defined as

$$F(t) = \int_0^t f(x)dx$$

Starting from these two functions, it is possible to define the s*urvivor function* $S(\cdot)$, which represents the probability that the failure event has not yet occurred by time *t,* and as such is of particular interest for risk analysis.

$$S(t) = \int_t^\infty f(x)\,dx = 1 - F(t)$$

In practice, the survival function can be expressed in several alternative ways; for simplicity's reasons we chose to use the nonparametric Kaplan–Meier estimator [9], which can be defined as

$$\hat{S}(t) = \prod_{i=0}^{t} 1 - \frac{d_i}{r_i}$$

where d_i is number of failures at time t_i, and r_i is the number of pipes at risk at time t_i (i.e., still functioning just before t_i).

Notice that these equations consider only the time component. Indeed, to handle covariates, such as the technology or other properties of the pipe, it is necessary to estimate separate functions (for example, one for each diameter). To choose which variables need to be taken into account, we performed various tests by estimating survivor functions with different groupings of the pipes using the features previously described. We assessed the resulting groupings through the log rank test [12], which is used to test the null hypothesis that there is no difference between the populations (in this case groups of pipes) in the probability of an event (in this case a failure) at any time point. We also verified the resulting risks on a test sample of pipes with the aid of domain experts. Interestingly, in many cases the log rank test suggested that there was a significant difference between the different groups (e.g., different technologies behaved very differently), however further analyses showed that the differences may be the result of biased and anomalous data. The result of this evaluation process led us to choose the hazard curves shown in Fig. 3 (left), which consider only the geographical risk. These are very simple models that can be easily interpreted and, as confirmed by domain experts, give reasonable insights to assess the probability of failure of the network's pipes. For comparison purposes, Fig. 3 (right) shows the hazard curve for all the recorded pipes.

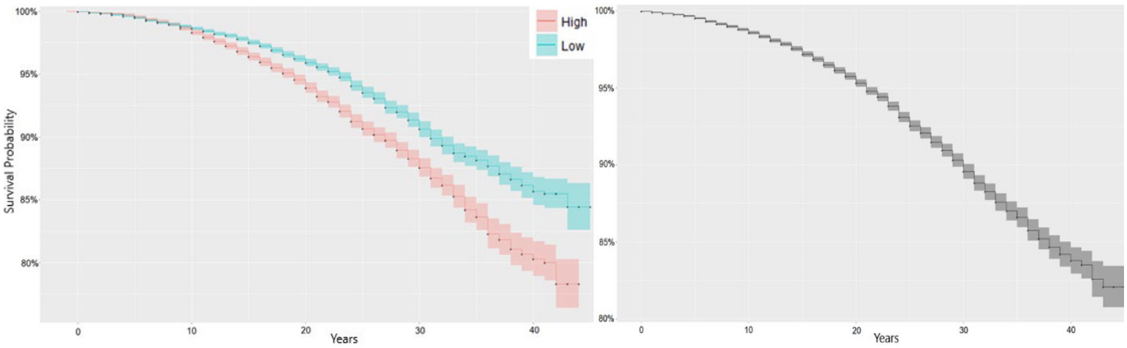

Fig. 3. Survival curve for the whole population of pipes (right) and for different types of geographical areas (left).

3.3. Maintenance optimization

In this section we formalize the mathematical model employed to define the optimal strategy. We want to plan the interventions on the network such that the maintenance cost is minimized. In particular, we plan the operations over a time horizon of Y_{plan} years such that the Net Present Value (NPV) of the maintenance cost is minimized over the next Y_{cost} years. Two types of elements constitute the network: pipes and manholes. A pipe $p \in P$ is characterized by its diameter d, length l, technology t and age a. A manhole $m \in M$ is uniquely characterized by a risk factor r, which is directly related to the number of years after which the manhole will require maintenance works, as they were conceived to allow interventions on the pipes, but are no longer required thank to modern technologies and should be removed whenever possible as they only represent a possible cost.

To state the decision support model, we need to declare all the costs involved in the process; *ACS* provided us with several tables of the costs involved, that we briefly describe and characterize in the following. A pipe $p \in P$ is associated with a maintenance cost $c_m(p)$, corresponding to the cost for repairing the pipe after a failure event, a substitution cost $c_s(p)$, required for its replacement with a new one, a malfunctioning cost $c_f(p)$, corresponding to the cost of leaking and other disservices following a failure event, and a reduced heat loss $c_h(p)$, due to the replacement of the pipe with a new one with better thermal insulation. Likewise, a manhole $m \in M$ is associated with a demolition cost $c_d(m)$ and a maintenance cost $c_n(m)$, directly related to its risk factor.

The optimization model has to decide *if* and *when* to replace a pipe or demolish a manhole, that is, plan the interventions over the time horizon Y_{plan}, taking into account the trade-off between the benefits connected to a newer pipe and its substitution cost: to do so, we compute for each year the corresponding total network's cost, that represents the sum of all the expected costs over pipes and manholes. For a pipe p and a cost term c, this is achieved by multiplying its failure probability on the year i, $\lambda^{(i)}(p)$, by the specified cost (see below for a detailed description). The same reasoning can be done to compute the costs associated with the set of manholes. It is clear that the costs computed in this way become a good statistical approximation of the true costs when the network size is big enough; a direct verification with *ACS* ensured our computed expected costs resulted close to their data.

The probability $\lambda^{(i)}(p)$ can be obtained from the results of the previous section by imposing the condition that the pipe breaks exactly between the year i and the following, given its age a. Denoting by T the failure event, we have:

$$\lambda^{(i)}(p) = P\left(T \geq a+i, T\langle a+i+1 |\ T\rangle a\right) = \frac{S(a+1) - S(a+i+1)}{S(a)}$$

In a similar fashion, the costs associated with a manhole are distributed linearly over the years, knowing that the risk factor is proportional to the number of years after which renovation works will eventually become necessary. After the first renovation, the manhole is assumed to have an expected remaining useful life of 25 years (and the costs are distributed accordingly).

A pipe replacement is represented by a boolean decision variable $x_p^{(i)} \in \{0, 1\}$, $i \in \{1, \ldots, Y_{plan}\}$, which is non-zero if the pipe p is replaced during the year i. Also, a manhole replacement is represented by a boolean decision variable $y_p^{(i)} \in \{0, 1\}$, with $i \in \{1, \ldots, Y_{plan}\}$.

We now have all the elements to formulate the expenses taken at each year $i \in \{1, \dots, Y_{cost}\}$. We separate each cost into two parts: the expenses due to non-replaced elements and those due to replaced ones. For the set of pipes and a generic cost term c, we can write:

- *Costs associated to non-replaced pipes*

$$\xi_r^{(i)}(c) = \sum_{p\in P} c(p) \cdot \lambda^{(i)}(p)\,(1 - \sum_{k=1}^{\min(i,Y_{plan})} x_p^{(k)})$$

- *Costs associated to replaced pipes*

$$\xi_{nr}^{(i)}(c) = \sum_{p\in P} \sum_{k=1}^{\min(i,Y_{plan})} c(p) \cdot \lambda^{(i-k)}(p)\, x_p^{(k)}$$

By substituting the cost c with the appropriate specific cost, we obtain all the needed terms: the total maintenance cost for each year for both replaced and non-replaced pipes, $C_{m,r}^{(i)} = \xi_r^{(i)}(c_m)$ and $C_{m,nr}^{(i)} = \xi_{nr}^{(i)}(c_m)$, the pipes' malfunctioning costs $C_{f,r}^{(i)} = \xi_r^{(i)}(c_f)$ and $C_{f,nr}^{(i)} = \xi_{nr}^{(i)}(c_f)$ and their reduced heat loss $C_{h,r}^{(i)} = \xi_r^{(i)}(c_h)$. In analogous way, by substituting $x_p^{(i)}$ with $y_m^{(i)}$, we obtain the maintenance cost for the non-demolished manholes $C_{n,r}^{(i)}$.

Finally, the total investment in year i correspond to the cost required for the substitution of pipes and the destruction of manholes,

$$C_s^{(i)} = \sum_{p\in P} c_s(p)\, x_p^{(i)} + \sum_{m\in M} c_s(m)\, y_m^{(i)}$$

The objective function to be minimized can be expressed as

$$\sum_{i=1}^{Y_{cost}} \frac{z^{(i)}}{(1+w)^i}, \quad z^{(i)} = C_s^{(i)} + C_{m,r}^{(i)} + C_{m,nr}^{(i)} + C_{n,r}^{(i)} + C_{f,r}^{(i)} + C_{f,nr}^{(i)} + C_{h,r}^{(i)}$$

where we assume $C_s^{(i)} = 0\, \forall i > Y_{plan}$ and w is the company's cost of capital.

We will not go into the detailed description of the imposed constraints, but rather list them below: some of them are structural, other were specifically required from *ACS*:

- A pipe is repaired at most once.
- A manhole can be demolished only when all the incident pipes belong to the most advanced technology (PR). When a manhole is demolished, it is removed from the network.
- There is a specific budget for retrofitting investments and another one for maintenance.
- Apart for the emergency interventions (maintenance), pipes' substitution has to be planned such that the involved pipe(s) exceed a minimum total length. This is to avoid 'spot' interventions and keep up with the fact that we did not include a thresholding cost to open a construction site, that instead exists.

The model obtained is an Integer Linear Programming model (ILP). The number of variables scales linearly with the number of pipes: this takes advantage of the network aggregation described in Section 3.1. The solver used was FICO Xpress, which was able to solve the network at hand (ca. 5k pipes and 2k manholes) in few minutes, with $Y_{plan} = 5$ and $Y_{cost} = 30$.

4. Results

We applied the methodology described in the previous sections to our use case, the DH network in Brescia, Italy. First, we validated the output of the risk analysis algorithm, simulating to be in 2012 and computing the risk index for the following years. Figure 5 shows boxplots for the results for years 2013 and 2016. The boxplots show the following information: the thicker line is the median value, the box represents the *interquartile range* (*IQR*) between the first (Q1) and the third quartile (Q3), while the lower and upper whiskers are respectively calculated as $Q1 - 1.5 * IQR$ and $Q3 + 1.5 * IQR$. The results indicate how the estimated risk would have been higher for those pipes that actually failed. In fact, the left (labeled as '1') and right column (labeled as '2') refer to the risk index distribution for those pipes that respectively did not and actually did fail, highlighting that the median values are

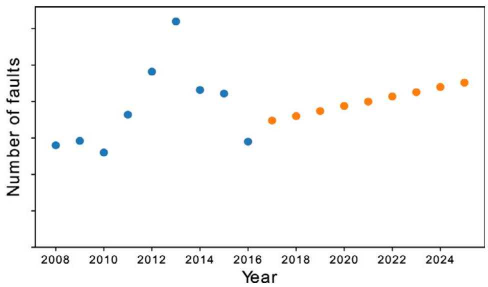

Fig. 4. Number of pipeline faults per year. Blue dots are the historical faults; orange dots are the expected number of faults according to the risk analysis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

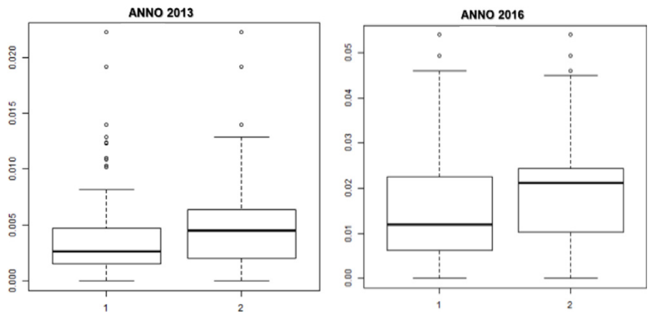

Fig. 5. Survival analysis validation on year 2013 and 2016 '1' is the distribution of the risk-index for non-faulty pipes; '2' is the distribution of the risk-index for faulty pipes.

markedly distinct even if with an overlap in the interquartile intervals. As expected, the confidence interval widens as we move further from the reference year. Moreover, the expected number of failures per year as of our risk index is in line with the trends of the previous years (see Figure 4).

Then, we used the model described in Section 3.3 to suggest optimal multi-year maintenance plans. The main goals of the study were to identify the most critical segments of the network, in terms of risk index and potential economic impact of an eventual failure, and to analyze how different drivers influenced the number of expected faults in the network, e.g., what would be the necessary budget to keep the number of faults steady through the years and whether that would be worth it from an economic standpoint.

Given a 5-year maintenance plan (i.e., the model schedules pipe replacement and manholes demolition activities in years 0 to 4, while only passive reaction to pipe failures is performed in subsequent years) and a time horizon upon which the *Net Present Value* (NPV) is evaluated as of 30 years, we benchmarked scenarios with different budgets, expressed in percentage with respect to the actual resource allocation by ACS in the reference year. The reference scenario was the so-called *Baseline*, representing the case in which no pipe replacement is foreseen.

The objective function of our model considers not only the investment cost of the piping replacement, but also the contribution of the potential inefficiencies caused by failures, which determine water and heat losses as well as operational disservices that may or may not overcome by regulation strategies. In general terms, a small number of pipe failures is desired, so that higher risk pipes should be addressed first, considering the economic implications of specific interventions.

The benchmark highlighted how the break-even point with respect to the baseline shifted forward as the budget increased, reflecting the higher initial capital costs. Yet, a higher budget (thus higher replacement ratio) ensured a reduction of the expected fault occurrences, with significant impacts not only from an economic, but also from a quality of service (and customer satisfaction) standpoint.

The smaller NPV of the maintenance costs evaluated over 30 years is linked to the BDG-100% scenario and the Baseline featured the highest figure, while the BDG-200% scenario performed halfway, yet with a 2.6% improvement with respect to the Baseline. The payback time was 15 years for BDG-100% and 25 years for BDG-200% (see Fig. 6).

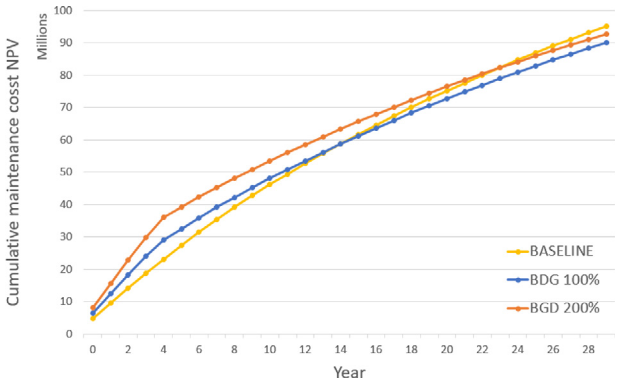

Fig. 6. Cumulative NPV of the maintenance costs for the baseline, the scenario with current budget (BDG-100%), and the scenario with doubled budget (BDG-200%).

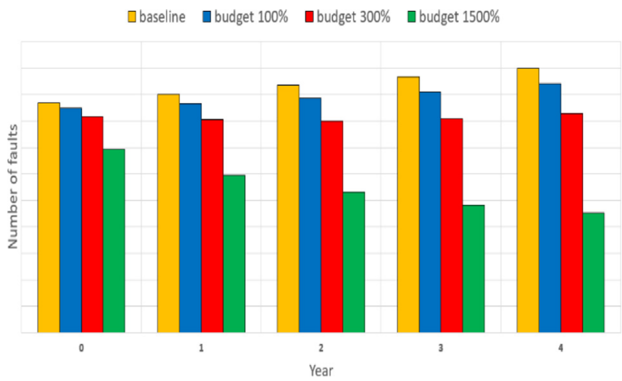

Fig. 7. Expected number of faults for different budget scenarios.

As shown in Fig. 7, the Baseline scenario (i.e., no replacement) outlined an increase of breakdowns by a factor of 20% in the timeframe considered, which could be mitigated by 10% with the current budget (BDG 100%) and be neutralized (i.e., steadying the number of failures) with a threefold increase in the set budget, while even higher figures (1500% of the current budget) would have been necessary to halve the number of failures in a 5-year span. Nevertheless, these scenarios proved to be not economically convenient, because the 30 year of horizon are not enough to recover such high investments. Moreover, a large budget also meant intervening in less significant areas, where the risk factors are lower, thus less impactful.

5. Conclusions

This work presented the maintenance planning optimization process applied to the DH network of Brescia, Italy. This is a real-life example of how a sustainable maintenance plan can be carried out by relying both on recorded data and on business strategies. The described approach starts from data analysis to objectively assess the risks of failures of the various components of the network, to incorporate these estimates in a more general optimization model that takes into account economic figures such as available budget, maintenance costs, and losses due to failures. By varying the available budget, it was possible to produce and analyze different maintenance scenarios, which could be considered in driving long-term strategic choices.

Several interesting directions for further work are available. The performed risk analysis could greatly benefit from a larger history of failures dataset, which would prevent many of the bias-related concerns, enabling adoption of more complex statistical and machine learning models. This said, the project represents a remarkable, real-life example of data- and model-driven analytics application to respond to the practical challenges of DH management, with a significant potential for replication.

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.

References

[1] Swanson Laura. Linking maintenance strategies to performance. Int J Prod Econ 2001;70(3):237–44.
[2] Ran Yongyi, et al. A survey of predictive maintenance: Systems, purposes and approaches. 2019, arXiv preprint arXiv:1912.07383.
[3] Sernhed Kerstin, Jönsson Mikael. Risk management for maintenance of district heating networks. Energy Procedia 2017;116:381–93.
[4] Khan Faisal I, Haddara Mahmoud M. Risk-based maintenance (RBM): a quantitative approach for maintenance/inspection scheduling and planning. J Loss Prev Process Ind 2003;16(6):561–73.
[5] Arzaghi Ehsan, et al. Risk-based maintenance planning of subsea pipelines through fatigue crack growth monitoring. Eng Fail Anal 2017;79:928–39.
[6] Tsitsifli Stavroula, Kanakoudis Vasilis, Bakouros Ioannis. Pipe networks risk assessment based on survival analysis. Water Resour Manage 2011;25(14):3729–46.
[7] Yeter B, Garbatov Y, Guedes Soares C. Risk-based maintenance planning of offshore wind turbine farms. Reliab Eng Syst Saf 2020;202:107062.
[8] de Almeida Costa Mariana, de Azevedo Peixoto Braga Joaquim Pedro, Ramos Andrade António. A data-driven maintenance policy for railway wheelset based on survival analysis and Markov decision process. Qual Reliab Eng Int 2021;37(1):176–98.
[9] Klein John P, Moeschberger Melvin L. Survival analysis: Techniques for censored and truncated data. Springer Science & Business Media; 2006.
[10] Christodoulou Symeon E. Water network assessment and reliability analysis by use of survival analysis. Water Resour Manage 2011;25(4):1229–38.
[11] Wang Ping, Li Yan, Reddy Chandan K. Machine learning for survival analysis: A survey. ACM Comput Surv 2019;51(6):1–36.
[12] Bland J Martin, Altman Douglas G. The logrank test. Bmj 2004;328(7447):1073.