Maciej ROMANIUK

ON SIMULATION OF MAINTENANCE COSTS FOR WATER DISTRIBUTION SYSTEM WITH FUZZY PARAMETERS

O SYMULOWANIU KOSZTÓW UTRZYMANIA DLA SYSTEMU DYSTRYBUCJI WODY O PARAMETRACH ROZMYTYCH*

In this paper we propose a model for evaluation of maintenance costs of a water distribution system (WDS). The set of possible states of each connection (i.e. a pipeline in the WDS) is related to various possible degrees of quality of the pipe and types of its malfunctions. The process of transitions between these states forms a semi-Markov process. Using Monte Carlo simulations, the length of services and the times of necessary replacements and repairs of the connections are obtained. These values are then used as an input for estimation of the maintenance costs of the whole WDS. During this step we take into account the concept of present value of money. Contrary to other approaches, instead of a constant yield, a stochastic process (the one-factor Vasicek model) of an interest rate is assumed. Then various simulated measures of reliability and the maintenance costs can be analysed, like an influence of various parameters of the pipes (e.g. intensities of damages) on the final costs of the performed services. They can be crucial in the analysis of risk for various possible decisions. Apart from the crisp approach, the Monte Carlo simulations are also applied, if some of the parameters of the WDS are fuzzified. Therefore uncertainty and experts' knowledge can be easily incorporated into the proposed procedure of the estimation of the maintenance costs. Observed differences between the crisp and the fuzzy output are highlighted. Simulation algorithms, necessary for both of these approaches, are also provided.

Keywords: water distribution system, maintenance costs, semi-Markov process, Monte Carlo simulations, present value.

W niniejszym artykule przedstawiono model obliczający koszty utrzymania i konserwacji dla systemu dystrybucji wody (WDS). Zbiór możliwych stanów każdego połączenia (tzn. odcinku rurociągu w WDS) jest zdefiniowany przez różne poziomy jakości rury oraz występujące typy uszkodzeń. Proces przejść pomiędzy tymi stanami jest opisany procesem semi-Markowa. Wykorzystując symulacje Monte Carlo, uzyskano długości okresów obsługi oraz momenty niezbędnych wymian i napraw. Wartości te są następnie wykorzystywane do estymowania kosztów utrzymania całego WDS. W kroku tym brana jest pod uwagę wartość pieniądza w czasie. W przeciwieństwie do innych podejść, zamiast stałej stopy procentowej, założono stochastyczny proces stopy (dany jednowymiarowym modelem Vasicka). Następnie na podstawie przeprowadzonych symulacji wykonano analizę opartą o różne miary niezawodności i obliczone koszty obsługi, np. zbadano wpływ parametrów połączenia (takich jak intensywność uszkodzeń) na ostateczne koszty konserwacji. Analizy tego typu mogą pełnić istotną rolę w ocenie ryzyka dla różnych możliwych do podjęcia decyzji. Poza podejściem typu crisp, zastosowano również symulacje Monte Carlo gdy niektóre z parametrów WDS zostały określone w sposób rozmyty. Dzięki temu można wykorzystać niepewność oraz wiedzę ekspercką w proponowanej metodzie estymacji kosztów obsługi. Zwrócono uwagę na różnice występujące pomiędzy podejściem crisp i rozmytym. Zostały również opisane niezbędne dla obydwu podejść odpowiednie algorytmy symulacyjne.

Słowa kluczowe: system dystrybucji wody, koszty konserwacji, proces semi-Markowa, symulacje Monte Carlo, wartość obecna.

1. Introduction

The main aim of each water distribution system (WDS, water pipes network) is to deliver water of desirable quality and assumed quantity for customers. In order to meet these requirements, maintenance services, like repairs and replacements of broken or malfunctioned parts, are required.

The literature devoted to the problems of simulation of the WDS behaviour and to analysis of its reliability is very rich (see, e.g, [3, 4, 7, 10, 11, 15, 17, 19, 25, 29] for a description of various approaches, and [16, 28] for a more detailed review of methods and literature). Because of a necessary long-time horizon planning (e.g. 20 or even 50 years), an influence of a value of the money in the future on maintenance costs should be taken into account. Also such a problem is addressed in many papers.

For example, in [20], the main aim is to estimate and validate various cost functions for different types of assets of a WDS, if their hydraulic (e.g., flow, pump head, pump power) and physical characteristics (e.g., volume, material, nominal diameter) are stored in a specially prepared data base. A method of a linear regression is used to model a dependency between various types of the costs (like an equipment cost) and the mentioned characteristics of a pipe (like a nominal diameter) based on data from several Portuguese urban water utilities. However, a method for a calculation of a present value of the total costs is not developed there in a more detailed way.

Contrary to the previous paper, in [12], let's say, a "macro-management" approach for a WDS rehabilitation problem with a real long-time horizon is widely discussed. In this paper, both an economics and hydraulic capacity of a WDS is analysed, if a deterioration of a pipe follows the Hazen-Williams equation. The total costs are related

to a rehabilitation of a pipe and some additional maintenance costs (mainly breakage repair costs). To model the breakage rate, the exponential dependency on an age of a pipe is assumed. The present value of the total costs is then calculated using a constant discount rate. The main aim of the whole procedure is to minimize the total costs associated with a rehabilitation, if some hydraulic constraints, necessary for a quality and a quantity of the supplied water, are preserved. As mentioned in this paper, an analysis period for such a procedure is typically equal to 30-60 years.

There are also more specialized approaches to the problem of the costs estimation. For example, in [25], an optimisation procedure for a problem of minimising the costs is also described. In this case, the authors express the total cost of a renewal, risk and an unavailability of a WDS as one function of time. But then this cost equation is minimized using the idea, that groups of pipes may be renewed together, if the benefits of grouping are balanced with the costs of shifting renewal investments in time. A special "greedy algorithm" is developed in order to find a solution for such a setup, and this numerical method is then applied for a WDS in a medium-sized Norwegian municipality. The authors informed about a high level of possible savings, if their "grouping" method is used.

Apart from planning a renewal schedule, the estimation of the costs can have impact on other types of decisions, even regarding physical characteristics of pipes. In [21], the costs related to a simple WDS are analysed for up to 50 years period. Two kinds of scenarios (with and without a replacement) are discussed, for which minimization procedures are applied using the installation and repair costs with, so called, a damage and inconvenience cost multiplier (which is related to the claims from damages and other inconveniences caused by breaks). In this case a genetic algorithm is used to solve the optimization problem. Then practical conclusions about the diameter of pipes, which should be installed, and an age, when instead of repairing, a replacement should be done, are drawn. Once again, the authors emphasizes possible savings, if an appropriate analysis of the problem is conducted.

The authors of [3] also focus on physical characteristics of a pipe. For various types of materials used for pipes, different models of break rates, based on an exponential hazard function, are incorporated. Such an approach leads to different functions of the costs, in which the costs and possible benefits of inspection technologies are also taken into account. It should be noted that in this paper, only a nominal value of the money is considered.

In some papers, time, which is necessary for a whole repair procedure, is completely neglected. Such an assumption may be accepted, even in practical situations (see, e.g., [28] for a more detailed discussion), but in many papers the required time is modelled using some random variable. For example, in [11], the generalized exponential density is applied. The authors also describe various types of the costs related to repairs, like the costs of a procedure itself, the costs of an extra energy (which is necessary to satisfy the public demands), the costs of water losses and a loss of revenues (which takes place, when the water demands are not satisfied).

A completely different scale (let's say, a "micro-management") of the problem of the costs is considered in [1]. The authors, using USA-CREL condition index (CI) system, optimize a maintenance time of a wastewater plumbing system for a single building – Esteghlal Hotel in Iran. The costs of repairs are given by a virtual variable related to a condition index of a component, and the costs of replacements are based on some physical characteristics of a pipe. In order to optimize the costs, a saving to investment method is applied. It allows the authors to analyse the maintenance management costs and it leads them to formulate practical advices concerning the time when the whole water system in the building should be replaced instead of extending its maintenance or additional repairs.

Then other approaches to a function of the costs are also possible. In [32], a reliability of sand filters, as a part of a whole WDS, is considered. For the costs function, a concept of Life Cycle Costing is adopted. Apart from various sources of possible payments (during a construction phase, a usage phase and, finally, a disposal), some other factors are taken into account: an economical operational readiness rate, a target operational readiness rate and a critical operational readiness rate. They form a whole concept of a risk analysis related to the possible costs.

As it is seen from these exemplary papers, the problem of an analysis and an estimation of the costs is very important during a construction and a further maintenance of a WDS. The approaches to such a problem are various, starting from a form of the cost function (which can depend on physical characteristics of a pipe, time of a repair etc.), through a scale of a WDS (a single building or a whole system), considered parts of a WDS (pipes, filters, intakes etc.), a scope of an analysis undertaken (which can include an additional optimization procedure, a calculation of some statistical measures etc.) and so on. However, we can point out some ideas, which are very common in these papers. Firstly, usually a long time period is considered, even 50-60 years. Secondly, some numerical algorithm is applied, which generates possible times of occurrences of the faults, varied values of the costs etc. Thirdly, this algorithm is related to some random process, which is assumed for the considered model. Fourthly, because of the mentioned length of the analysed period, the value of the money in time is usually taken into account.

Such an approach (i.e. introduction of the relation between the time and the value of the cash flows) is widely studied in financial mathematics, because it is known, that one unit of money, which is paid now, has different value than the same unit, which will be paid in the future. Therefore, we also adopt this idea and analyse the estimated present value (i.e. the value of money for now) of the costs of future services. The maintenance costs are one of the key factors during selection from a set of various possible decisions, if the financial risk is taken into account. However, in various papers, the related discounting factor is only a minor, almost negligible, part of the whole simulated model. Usually, the yield is given by a factor, which is constant in time. In our considerations, the stochastic model of the interest rate is directly embedded into the Monte Carlo simulations, as important part of a reliability analysis.

Additionally, it is given by the one-factor Vasicek model. Therefore, such a yield can be better adjusted to real life data (e.g. estimated using statistical methods), especially if a long time horizon is considered. Also other, more complex models of the interest rate process can be directly used in the general approach presented in this paper. To the best knowledge of the author, an incorporation of a stochastic model of the yield is not considered in other papers concerning the analysis of the WDS costs at all. Further, using the Monte Carlo simulations, we show that there are significant differences in the estimated output, if, instead of the stochastic model, more classical approaches (like a constant yield) are taken into account. It means that, if we ignore a stochastic nature of the yield (which, of course, can not be constant all the time), some significant error is introduced into our analysis of the costs.

Additionally, in many of the papers, only one source of an uncertainty is modelled, i.e., some random process describes a behaviour of a WDS, e.g. the times of its defaults. However, there are other possible sources of uncertainties, which can be related to an incorporation of the experts' knowledge. Therefore, in our setting, the parameters of the model are described by fuzzy numbers. It means, that they are not completely precise (i.e., "crisp") but they are, in some way, uncertain ("near to / about"). Such an idea is close to the real life, because, e.g., an unconditional time of a replacement of a pipe is rather "about 5 years", than strictly "always and only 5 years and not one day more / less".

In this paper, the costs of the services are related to a set of possible states of the single connection. These states reflect various types of quality and degradation of the pipeline and its current behaviour (i.e. if it operates). We assume that the random transitions between the states of the single connection form a semi-Markov process. These transitions are also influenced by a deterministic, unconditional replacement age, which can be part of a maintenance routine.

The connections in the WDS can be grouped by their types, depending on various technical properties (like quality, size etc.), which leads to different sets of statistical and reliability parameters. Then, the present value of the future possible costs of the maintenance services is estimated using Monte Carlo simulations for the whole WDS, which consists of the mentioned groups. Apart from the estimator of the present value, other statistical measures related to the repairs and the replacements, as well as their costs, can be also directly obtained using simulations. These measures can be an important input for the reliability analysis of the WDS.

This paper can be seen as an important step in generalization of the approach proposed in [26]. The currently presented contribution is fourfold. Firstly, as it was previously mentioned, instead of the constant yield, the interest rate is modelled using a stochastic process. The Monte Carlo simulations of trajectories of this process are directly embedded into a procedure of generation of the costs of the services of the whole WDS. The stochastic model allows us for a better adjustment of the yield to the real life data and to minimise the possible error related to an assumption of the constant yield. This is very important for the practitioners in a proper estimation of the future maintenance costs. Secondly, instead of a classical Markov process with a discrete state space, the semi-Markov process of the transitions between the states, with the introduction of the deterministic, unconditional replacement age, is proposed. It allows us to apply other random distributions for the transitions, not only the exponential one. Thirdly, we enrich our model introducing fuzzification of some parameters of each connection in the WDS, like the costs of the services and the intensities of the transitions. This fuzzification reflects the uncertainty observed in real life data, which can be overcome with the experts' knowledge. It allows us, to some extent, to measure "an error" produced by these uncertain parameters, because a simulated output forms also a fuzzy number. This is important for a maintenance analysis, because it leads to a better estimation of the future, uncertain costs. Fourthly, apart from the Monte Carlo simulations of the WDS for the crisp case, some examples of the analysis of the WDS, if the parameters are given by fuzzy triangular numbers, are presented. It should be mentioned, that other types of L-R fuzzy numbers can be also directly applied for the considered approach.

This paper is organized as follows. In Section 2, the model of the WDS is proposed. In this part, the set of connections, the model of the maintenance costs related to repairs and replacements, as well as the one-factor Vasicek model are presented. In Section 3, a Monte Carlo approach is used to simulate a behaviour of the WDS, if all of the considered parameters are crisp. Apart from the description of the relevant algorithm, some examples of various settings are analysed. Section 4 is devoted to the problem of fuzzification of the parameters of the model. Firstly, some definitions and necessary symbols are introduced. Then, various examples of the output for the fuzzy environment, which are based also on the Monte Carlo simulations, are discussed.

2. Model of WDS

Let us suppose that the considered WDS is modelled by a graph of connections G. In this graph, each connection (i.e. a pipeline which is a part of the whole WDS) is represented as an edge, and possible sources or outflows are denoted by nodes. In practice, various types of

nodes are possible, e.g., intake points (sources), supply points (sinks), branching nodes (see, e.g., [18] for a more detailed discussion).

2.1. States of the connection

In the following, we focus only on the edges of the graph G, i.e. the connections of the considered WDS. We assume that the state of each connection for a given time t can be described by one of the possible states from a set $S=\{0,1,2,3,4\}$. These numbered states are related to the various types of quality of this pipeline (which depends on its degradation and the probability of possible malfunction) and its current behaviour (i.e., if this connection operates or not). In this way we have:

– **State 0** The considered section of WDS is *under replacement* now, because it is completely broken, so it does not work.
– **State 1** The section is *under repair*, because of its temporary malfunction, and it can not be used now.
– **State 2** The section works and can be described as to be in a *burn-in phase*, when some initial defects just after starting this part (as a new one) are possible.
– **State 3** The section operates and it is in its *normal operations state* (or a *standard state*), because the previous period of occurrence of some initial failures has ended.
– **State 4** The section works, but a *wear-out period* after the normal operations state has been reached, so this part of the WDS may fail with higher probability due to deteriorating stresses from its previous lifetime.

Some of the considered states, namely burn-in, normal operations and wear-out phases, are described in a more detailed way in [4]. The quality of a pipeline can be also modelled using other approaches, see, e.g., [7, 17, 25, 28, 29]. However, the presented setting is more flexible, because it allows for an introduction of more than the five considered states. Therefore, the quality of a pipeline can be modelled even in a more complex way, if it is necessary for some practical purposes. Additionally, the parameters of each state can be statistically estimated independently of other states.

Our main aim is to analyse the state of the j-th connection at the time t, which is defined by a *process of the state of the connection* $S^{(j)}(t)$. From our previous assumption, we get $S^{(j)}(t) \in \{0,1,2,3,4\}$, i.e. for any time t, the state of the j-th connection is described by some state from the set S.

Then, $S^{(j)}(0)$ is a starting state for the j-th connection, i.e., the state of the pipe during initialization of the whole algorithm (or estimation procedure, equivalently). The behaviour of the process $S^{(j)}(t)$ for the fixed j is described by random sequences $L_1^{(j)}, L_2^{(j)}, \ldots$ and $S_1^{(j)}, S_2^{(j)}, \ldots$. Variables in the first sequence are periods between the moments when the j-th connection enters into some state and then leaves this state in order to proceed to another state. The second sequence describes the consecutive states for the j-th connection at the moments $L_1^{(j)}, L_1^{(j)} + L_2^{(j)}, \ldots$.

We assume that the process formed by $\left(S_i^{(j)}, L_1^{(j)} + \ldots + L_i^{(j)} \right)_i$ is a Markov renewal process. Then, straightforwardly,

$$N^{(j)}(t) = \max_n \left\{ L_1^{(j)} + \ldots + L_n^{(j)} \leq t \right\}$$

is a Markov renewal counting process (i.e., the number of the transitions between the states in our case) and $S^{(j)}(t) = S_{N^{(j)}(t)}^{(j)}$ is a semi-Markov process.

Additionally, we assume that for every i,j the processes $S^{(i)}(t)$ and $S^{(j)}(t)$ are independent of each other in probabilistic sense. From

the practical point of view, it means that there is no "information flow" between the connections and that the state of one part does not depend on the condition of other pipeline.

If for any j, $L_1^{(j)}, L_2^{(j)}, \ldots$ are *iid* random variables distributed according to the exponential distribution, then the process $S^{(j)}(t)$ directly reduces to a Markov chain with continuous time t and the discrete finite state space S (see, e.g., [22]). However, due to next assumptions and incorporation of the fuzzy approach, such a simplification is not relevant for the case discussed in this paper.

Let $R^{(j)}$ denotes the deterministic and unconditional replacement age (i.e. the planned replacement). It means that when the length of period after the last replacement of the j-th connection surpasses $R^{(j)}$, then this pipeline is always replaced regardless of its previous history (see also [19]). Mathematically, we calculate:

$$R_*^{(j)} = \max_{0 \le t_1 \le t_2} \left\{ t_2 - t_1 : S^{(j)}(t_1) = 0, S^{(j)}(t) \ne 0 \; for \; t \in (t_1, t_2] \right\} \quad (1)$$

and when $R_*^{(j)} \ge R^{(j)}$, the process $S^{(j)}(t)$ immediately changes its value to the state 0 (i.e. under replacement).

It is easily seen, that if the deterministic transitions given by $R^{(j)}$ are not taken into account, the behaviour of the process $S^{(j)}(t)$ is completely described by the Markov renewal process $\left(S_i^{(j)}, L_1^{(j)} + \ldots + L_i^{(j)} \right)$. In order to conduct the simulations, the transition probabilities:

$$Q_{kl}^{(j)}(t) = Pr\left(S_{n+1}^{(j)}(t) = l, L_{n+1}^{(j)} \le t \mid S_n^{(j)}(t) = k \right) \quad (2)$$

i.e. the distribution of the time t of the transition from the state k to the state l for each connection j, should be set. In the following we assume that:

$$Q_{kl}^{(j)} = \min_{l \in S} \left\{ X_{kl}^{(j)} \right\}, \quad (3)$$

where $X_{k0}^{(j)}, \ldots, X_{k4}^{(j)}$ are independent random variables and each $X_{kl}^{(j)}$ is given by the exponential distribution with the intensity $\lambda_{kl}^{(j)}$, i.e. $X_{kl} \sim Exp\left(\lambda_{kl}^{(j)} \right)$. It means that the distribution of $Q_{kl}^{(j)}$ is given by the minimum of random times of transitions from the state k to each other possible state. In the case of the exponential distribution, the formula (3) is consistent with a classical stochastic approach, i.e. with the previously mentioned Markov chain. However, in practical applications other kinds of distributions could be also used to model the random variables $X_{k0}^{(j)}, \ldots, X_{k4}^{(j)}$, like, e.g., some lognormal distribution. Especially, the formula (3) leads directly to the method which can be applied in numerical simulations of the transitions between the states even in the fuzzy case considered in Section 4.

If we restrict ourselves to the exponential distributions for $\left\{ X_{kl}^{(j)} \right\}_{k,l \in S}$ for the fixed connection j, then we have an *intensity matrix*:

$$\Lambda^{(j)} = \left(\lambda_{kl}^{(j)} \right)_{k,l \in S}$$

In practice, the values in $\Lambda^{(j)}$ should be estimated from real-life observations. In order to do this, we should note, that a mean sojourn time for the state k is given by:

$$\tau_k^{(j)} = \frac{1}{\sum_{l \ne k} \lambda_{kl}^{(j)}} \quad (4)$$

and a one-step probability of the transition between the states k and l is equal to:

$$Pr_{kl}^{(j)} = Q_{kl}^{(j)}(\infty) = \frac{\lambda_{kl}^{(j)}}{\sum_{m \ne k} \lambda_{km}^{(j)}} \quad (5)$$

Then, using a set of equations based on (4) and (5) for all of the states, it is possible to set values of the intensities in the intensity matrix calibrated to our observations. For example, $\tau_1^{(j)}$ is a mean time for a repairing procedure, $\tau_2^{(j)}$ is a mean time, when a pipe is in its burn-in phase, $Pr_{12}^{(j)}$ is a probability, that a pipe is repaired in such a way, that it can be treated to be in the burn-in phase afterwards etc. Of course, this procedure may require some additional experts' knowledge (e.g. about a mean time for the burn-in phase). However, this requirement directly coincidences with an introduction of a fuzzy approach described in Section 4.

If, instead of the exponential distributions, other kinds of densities are considered, the equations similar to (4) and (5) can be also used. We have (see, e.g., [13])

$$\tau_k^{(j)} = \int_0^\infty \prod_{l \ne k} \left(1 - F_{kl}^{(j)}(t) \right) dt \, ,$$

where $F_{kl}^{(j)}(t)$ is a cdf of the time spent in the state k and associated with the transition from k to l, and

$$Q_{kl}^{(j)}(t) = \int_0^t \prod_{m \ne k} \left(1 - F_{km}^{(j)}(x) \right) dF_{kl}^{(j)}(x)$$

for a relevant kernel matrix. These integrals can be analytically intractable for more complex distributions, but it is still possible to calibrate the necessary parameters for the assumed densities even in practical applications (see, e.g., [13] for a more detailed discussion and some examples).

As it was previously mentioned, in the case of the set S considered in this paper, only some transitions between the states are possible. Therefore, only some intensities $\lambda_{kl}^{(j)}$, related to the exponential distributions are not equal to zero. So, the intensity matrix has the form

$$\Lambda^{(j)} = \begin{pmatrix} 0 & 0 & X & 0 & 0 \\ 0 & 0 & X & X & X \\ X & X & 0 & X & 0 \\ X & X & 0 & 0 & X \\ X & X & 0 & 0 & 0 \end{pmatrix} \quad (6)$$

where X denotes a non-zero entry. Rows and columns of the matrix (6) are numbered in the same way as the states from the set S, i.e. from 0 up to 4. For example, $\lambda_{02}^{(j)}$ corresponds to the intensity of replace-

ment of the broken part by a new one, which begins its work in the burn-in phase afterwards, and $\lambda_{30}^{(j)}$ describes the intensity of breaking down with a necessary replacement as its consequence, if this connection was in its normal operations state previously.

Additionally, from the description of the states, we assume that:

$$\lambda_{20}^{(j)} > \lambda_{30}^{(j)}, \lambda_{40}^{(j)} > \lambda_{30}^{(j)}$$

and:

$$\lambda_{21}^{(j)} > \lambda_{31}^{(j)}, \lambda_{41}^{(j)} > \lambda_{31}^{(j)}$$

These inequalities reflect the fact that the burn-in and wear-out states have greater intensities of both replacements and repairs than the normal operations phase, as stated previously.

During the Monte Carlo simulations considered in this paper, the process of the state of the connection $S^{(j)}(t)$ is, in addition, modified by the instantaneous transitions caused by the relevant value of $R^{(j)}$.

In the following we assume that all of the connections could be grouped in K groups by their types which depend on various technical data for the pipes, e.g., their dimensions, a kind of material which is used etc. (see, e.g., [3] for additional details). Straightforwardly, these types are related to different intensity matrices and, e.g., values of the starting states and unconditional replacement ages.

2.2. Model of maintenance costs

Based on the model described in Section 2.1, the Monte Carlo approach can be applied to simulate a trajectory of the process $S^{(j)}(t)$ for each connection j, and then, in a similar way, to generate a behaviour of the whole WDS. The lengths of the periods between the transitions are part of such an output and the times of entries to the states 0 or 1 (i.e. the necessity of replacement / repair of some part) can be directly calculated. In the following, these times are denoted by $t_1, t_2, \ldots$.

In this paper we focus only on the maintenance costs related to replacements and repairs. Of course, other types of costs (like costs of water losses, loss of revenues etc. – see, e.g., [4]) can be also incorporated into the considered numerical model.

We assume that the mentioned costs depend on the type of maintenance service (i.e., if replacement or repair of some part is necessary), the length of this service and the type of the connection. Therefore we have:

$$\mathrm{c}^{(j)}(t_i) = c_{const}^{(j)} + c_{var}^{(j)} \tag{7}$$

where $\mathrm{c}^{(j)}(t_i)$ denotes a *total sum of costs* for the given j-th connection and the moment t_i of starting the necessary service, $c_{const}^{(j)}$ is some constant value independent of the length of the maintenance, i.e. it is a fixed cost, and $c_{var}^{(j)}$ denotes a variable cost, i.e. a value which depends on the length of this service.

It is possible to apply various approaches for calculation of the variable costs using the Monte Carlo simulations. For example, these costs may be directly related to the length of period of the labour (and then strictly deterministic), or modelled by an additional random distribution which depends on the mentioned length, the type of connection and other parameters (which directly leads to a random value of the payment).

2.3. Model of interest rates

In the considered setting we assume that the value of money depends on time, i.e. we apply the concept of the present value, which is widely known in financial mathematics (see, e.g. [5, 27]). Such a set-up is especially useful if we are interested in the long time horizon T (e.g., 10–20 years) for which the estimated costs of the maintenance services should be found. Then, the values evaluated for different scenarios can be compared for now (i.e. when $t=0$). These scenarios can be directly related to various possible decisions (e.g., different values of $R^{(j)}$), so they lead to a selection of the best decision if the financial risk is taken into account.

In the following, we evaluate the present value of the total sum of the costs of repairs and replacements $PV(c)$, which is given by:

$$PV(c) = \sum_{i,j} PV\left(\mathrm{c}^{(j)}(t_i)\right) \tag{8}$$

In order to calculate (8), the relevant model of the interest rate should be used to find the discounting factor $PV(.)$ for each $\mathrm{c}^{(j)}(t_i)$. There are many such models known in financial mathematics (see, e.g., [5]), but in this paper we focus on the one-factor Vasicek model described by the formula:

$$dr_t = a(b - r_t) + \sigma dW_t, \tag{9}$$

where r_t is a value of the interest rate at time t, W_t is the standard Brownian motion, and a, b, σ are parameters of this model. Moreover, b characterizes a long term mean level (i.e. the trajectory of r_t is directed to this value during its long run), a reflects a speed of reversion towards b, and σ is an instantaneous volatility (variability) of the trajectory introduced by the random component of the Brownian motion.

In our setting, the model of the interest rate is directly embedded into the Monte Carlo simulations, as explained in Section 3. Therefore, it is necessary to apply the relevant iterative scheme for generation of increments Δr_t and the discounting factor. In the case of the model (9), such a scheme (see, e.g., [5] for more details) is based on evaluation of r_t at the fixed moments $0 = t_0 < t_1 < \ldots < t_n$, using the formula:

$$r_{t_{i+1}} = exp\left(-a(t_{i+1} - t_i)\right)r_{t_i} + b\left(1 - exp\left(-a(t_{i+1} - t_i)\right)\right) + \sigma\sqrt{\frac{1 - exp\left(-2a(t_{i+1} - t_i)\right)}{2a}}Z_i,$$

where $Z_1, Z_2, \ldots, Z_n$ are *iid* samples from $N(0,1)$. Then, the cdf of $r_{t_{i+1}}$ for the given value of r_{t_i} is equal to:

$$r_{t_{i+1}} \sim N\left(exp\left(-a(t_{i+1} - t_i)\right)r_{t_i} + b\left(1 - exp\left(-a(t_{i+1} - t_i)\right)\right), \sigma^2\frac{1 - exp\left(-2a(t_{i+1} - t_i)\right)}{2a}\right) \tag{10}$$

In the same way, the factor:

$$fv_{(t_i, t_{i+1})} = \int_{t_i}^{t_{i+1}} r_s ds,$$

which is necessary to evaluate the present value, can be found. The cdf of $fv_{(0, t_{i+1})}$ for the given $fv_{(0, t_i)}$ and r_{t_i} (see, e.g., [5]) is equal to:

$$fv_{(0,t_{i+1})} \sim N\left(fv_{(0,t_i)} + \frac{1}{a}\left(1 - exp\left(-a\left(t_{i+1} - t_i\right)\right)\right)r_{t_i} + \frac{b}{a}\left(exp\left(-a\left(t_{i+1} - t_i\right)\right) + a\left(t_{i+1} - t_i\right) - 1\right), \frac{\sigma^2}{a^2}\left(\left(t_{i+1} - t_i\right) + \right.\right.$$
$$\left.\left. \frac{1}{2a}\left(1 - exp\left(-2a\left(t_{i+1} - t_i\right)\right)\right) + \frac{2}{a}\left(exp\left(-a\left(t_{i+1} - t_i\right)\right) - 1\right)\right)\right).$$

(11)

The conditional covariance inside the pair $\left(r_{t_{i+1}}, fv_{(0,t_{i+1})}\right)$ for the fixed value of $\left(r_{t_i}, fv_{(0,t_i)}\right)$ is equal to:

$$\frac{\sigma^2}{2a}\left(1 + exp\left(-2a\left(t_{i+1} - t_i\right)\right) - 2exp\left(-a\left(t_{i+1} - t_i\right)\right)\right).$$

2.4. Parameters of the model

Taking into account our previous considerations, the parameters of the whole model can be divided into two groups:
1. Parameters of the given type $k=1,...,K$ of the connection related to its reliability and its maintenance costs: $S^{(k)}(0)$ – the starting state of the connection, $R^{(k)}$ – the unconditional replacement age, $\Lambda^{(k)}$ – the intensity matrix, $c_{const}^{(k)}$ – the fixed costs ($c_{const,rpl}^{(k)}$ for the replacement and $c_{const,rpr}^{(k)}$ for the repair), $c_{var}^{(k)}$ – the variable costs ($c_{var,rpl}^{(k)}$ – for the replacement, $c_{var,rpr}^{(k)}$ – for the repair), n_k – the number of the connections of this type.
2. Parameters of the interest rate model related to the financial setup: a – the speed of reversion, b – the long term mean level, σ – the instantaneous volatility.

3. Simulations – the crisp case

In this part we assume that all of the parameters mentioned in Section 2.4 are given by crisp values (i.e. real numbers). Because of this assumption, the relevant algorithm for simulation of the behaviour of the WDS and the interest rates is more straightforward than in the fuzzy case considered in Section 4.

3.1. Algorithm

In order to calculate the present value of the maintenance costs and other output which is useful for the reliability and maintenance analysis (see Section 3.2 for some examples), the three main phases of the algorithm should be repeated n times, where n is an overall number of simulations (see Appendix, Algorithm 1).

During the first phase, the relevant Markov renewal process is simulated using the random transition probabilities given by (2) (in the general case) or (3) (for the exponential distribution considered further on in this paper). Additionally, the condition (1) should be checked and, if $R_*^{(j)} \geq R^{(j)}$, the state of the connection is deterministically set to 0. Then, all of the transitions to the states 0 or 1 for each trajectory are found and the relevant times and periods of the maintenance services are calculated for these events. It leads us to an evaluation of the nominal total sums of the costs for each generated time of the necessary service using the formula (7).

In the second phase, the iterative schemes (10) for r_t and (11) for $fv_{(0,t)}$ are used to generate the trajectory of the interest rate process (9), and, simultaneously, the discounting factor. During this Monte Carlo step, the ordered times of the

transitions to the states 0 or 1 of all of the trajectories simulated in the first phase are taken into account.

In the third phase, the estimator of the total discounted costs of the maintenance services is calculated.

3.2. Examples of analysis

In each of the following examples $n = 1\,000\,000$ simulations are conducted.

3.2.1. Example I

In the following, for illustration purposes, i.e. to validate a correctness of the introduced algorithms and to present possible outcomes, we use artificial parameters for the considered WDS. However, some of them were obtained from an oral communication with the experts. Let us assume that only one type of the pipeline is considered and it is described by the parameters:

$$S^{(1)}(0) = 2, R^{(1)} = 5, c_{const,rpl}^{(1)} = 5, c_{const,rpr}^{(1)} = 4, c_{var,rpl}^{(1)} = 3, c_{var,rpr}^{(1)} = 2, n_1 = 20$$

(12)

Additionally, the intensity matrix of the transitions is given by:

$$\Lambda^{(1)} = \begin{pmatrix} 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 24 & 26 & 0 \\ 0.5 & 1 & 0 & 1 & 0 \\ 0.4 & 0.9 & 0 & 0 & 0.3 \\ 0.6 & 1.1 & 0 & 0 & 0 \end{pmatrix}$$

(13)

and the unit of time can be identified with one year. Therefore an average time which is necessary for complete replacement of the broken connection is about one month, an average time of repair (if then the burn-in phase is achieved) is half of this time, the unconditional replacement age is equal to five years etc. Of course, in practical applications the relevant parameters of the considered connections should be estimated from real data (or based on the experts' opinions), as noted in Section 2.1.

For the interest rate model it is assumed that:

$$a=0.1, b=0.05, r_0=0.04, \sigma=0.001$$

(14)

and we are interested in a rather long twenty years time horizon of the financial analysis of the maintenance costs (i.e., T=20).

After using the Monte Carlo simulations, the following estimators are found: $\bar{x}_{rpr} = 360.597$ (an average number of repairs), $\bar{x}_{rpl} = 190.164$ (an average number of replacements),

Table 1. *Exemplary estimators from Example I*

	Minimum	Q1	Mean	Q3
Times of repairs	5.58913e-11	0.00574441	0.0199603	0.0276658
Times of replacements	3.55271e-15	0.0238704	0.0829755	0.115023
Costs of repairs	4	4.01149	4.03992	4.05533
Costs of replacements	5	5.07161	5.24893	5.34507
	Maximum		Stand. dev.	
Times of repairs	0.333829		0.0199602	
Times of replacements	1.64975		0.0829587	
Costs of repairs	4.66766		0.0399203	
Costs of replacements	9.94925		0.248876	

$\bar{x}_{rpl,R} = 18.1029$ (an average number of the unconditional, planned replacements), $\bar{x}_{rpl,NR} = 172.061$ (an average number of the replacements without the planned ones, i.e. the not planned replacements) and $PV(c) = 1629.2$ (the discounted value of the maintenance costs) for the whole considered period. Some other useful estimators of times and costs can be also directly found from the same sample (see Table 1 for some examples).

The obtained estimator of the discounted value of the maintenance costs for the Vasicek model $PV(c)$ can be directly compared with the similar values for more classical approaches – a nominal value of the cash flow (denoted further by $PV_{nomin}(c)$) and the model with the constant yield r (denoted by $PV_{const}(c)$). Using the respectively modified Monte Carlo approach, we get $PV_{nomin}(c) = 2454.9$ and $PV_{const}(c) = 1547.42$. In this second model, $r = 0.05$ is set, which is equal to the long term mean level b for the Vasicek model given by (14). The relative differences between the outputs from these models are rather significant – about 50.6813%, if we compare the nominal approach with the Vasicek model, and about −5.01964%, if the constant yield is taken into account. Therefore, a selection of the appropriate interest rate model seems to be an important step in a decision making process.

3.2.2. Example II

The considered Monte Carlo approach could be also useful for decision makers to examine an influence of various parameters or scenarios on the estimated output like the costs of maintenance services or the reliability statistics. For example, from the previous analysis the average number of the unconditional replacements may be seen as too high, so one may be tempted to increase the unconditional replacement age and set the higher value of $R^{(1)}$. Of course, such an action could have undesirable effect in the long-time horizon and, if $R^{(1)}$ is set too high, then also the number of "usual" (i.e. not planned) replacements or repairs caused by deterioration processes may increase. But the simulations, similar to the previously described ones, can be directly used to support the process of taking appropriate decisions.

For example, for the parameters considered in our setting, such a course is desirable, because for $R^{(1)} = 10$ we have $\bar{x}_{rpr} = 361.715$, which is similar to the value from the previous case (but now this average is also higher, the relative difference is about 0.310041%), $\bar{x}_{rpl} = 173.847$ is significantly lower value, $\bar{x}_{rpl,R} = 1.2931$ is only 7.14% of the previous average of number of the unconditional replacements, $\bar{x}_{rpl,NR} = 172.554$ (which is similar to the previous case, but once again the relative difference is positive and equal to 0.286468%), and the total discounted value of the cash flow is also reduced to $PV(c) = 1578.87$.

A similar analysis can be done also for a whole interval of the possible values of $R^{(1)}$. Figure 1 presents the relation between the unconditional replacement age and $PV(c)$, in the similar way as in Example I, for different models of the interest rate: the Vasicek model (circles), the constant rate (squares) and the nominal value (rhombuses). All of these present values are exponentially decreasing functions, but the differences between them are, once again, significant.

The similar analysis can be done also for other values, which are very important for the practitioners and the decision makers, like the average number of repairs $\bar{x}_{rpr}$ (see Figure 2) and the average number of not planned replacements $\bar{x}_{rpl,NR}$ (see Figure 3). In the both cases,
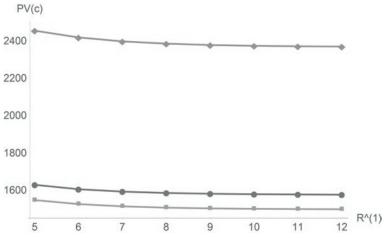

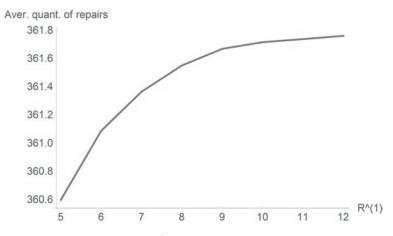Fig. 1. Graph of relation between $R^{(1)}$ and $PV(c)$ from Example II

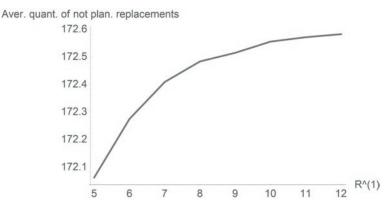Fig. 2. Graph of relation between $R^{(1)}$ and $\bar{x}_{rpr}$ from Example II

Fig. 3. Graph of relation between $R^{(1)}$ and $\bar{x}_{rpl,NR}$ from Example II

the average number of services is a slowly increasing function of $R^{(1)}$ for the given set of the parameters.

4. Fuzzification of the parameters

As it is known, some sources of uncertainty may be easily modelled by the fuzzy approach (see, e.g., [6]). In such a case the value of some uncertain parameter is based on expert's knowledge. This approach is especially very important when data is sparse and various data analysis methods, like statistics, are not usable or even not

possible. Then, based on opinions of the experts, the necessary parameters of the model can be evaluated. However, usually these opinions have not completely precise form (e.g. like real numbers), they are rather stated as "the value of this parameter is about 5", so they can be transformed into special form, known as fuzzy numbers. The fuzzy numbers are also used during making statistical decisions (see, e.g., [9] for an example of application to some reliability data).

In the following, we assume that some of the parameters describing the considered type of the connection (enlisted in Section 2.4) are fuzzified. Such an assumption requires different Monte Carlo approach than in the crisp case, therefore new, relevant algorithm is discussed in Section 4.2. A comparison of the Monte Carlo simulations of crisp and fuzzy random numbers can be found, e.g., in [8].

For example, the planned replacement age $R^{(j)}$ is treated further as a fuzzy number. Theoretically, such a parameter is completely deterministic and crisp value, e.g. we may say "$R^{(1)}$ is equal to 5 years". However, in the practical cases, this value is not completely exact and crisp and it is rather "about 5 years plus / minus a few days / months", which is caused by a temporary lack of funds for the given period, urgent and more important repairs in other areas just in the given moment (hence, lack of the experienced staff at that time), problems with preparing plans for the traffic exactly in the given date (like busy working day) etc. Therefore, the introduced fuzziness has practical basis.

However, some quantitative information about the planned replacement age or other parameter is still required. Let us assume, that the experts express their opinions. These sentences can have, e.g., a form of linguistic variables (like "a time of a planned replacement is rather later than sooner"). But in the setting considered in this paper, it is necessary to "translate" them into relevant sets of real numbers with membership functions, i.e. fuzzy numbers (see, e.g., [30]). Then, these numbers, which have to satisfy some additional requirements about a monotonicity of their membership functions (like L-R numbers), are used during simulations in order to evaluate a fuzzy output, e.g., fuzzy maintenance costs. Applying such a procedure, we gain an additional knowledge about a dependency between the uncertain (fuzzy) parameters (i.e. the experts' opinions) and a simulated (uncertain) outcome (i.e. conclusions, which are important for our management decisions). It means, that a level of the uncertainty in the input is directly reflected in the uncertainty of the output and can be also easily measured. For example, a difference between the fuzzy maintenance costs for two scenarios – "the planned replacement age is equal to 5 years plus / minus a half of the year" and "the planned replacement age is equal to 4.5 years plus / minus a half of the year" – can be exactly seen and measured, and a relevant decision can be taken.

It should be also noted, that an incorporation of the fuzzy variables is related to a different type of uncertainty, than one modelled by probabilistic approaches. Then, we enrich our model in a significant way.

4.1. Preliminaries

Now we present basic definitions and notation concerning the fuzzy approach, which will be used in the further part of the paper. Additional details can be found in, e.g., [14].

For a fuzzy subset $\tilde{A}$ of the set of real numbers R we denote by $\mu_{\tilde{A}}$ its membership function $\mu_{\tilde{A}}: R \rightarrow [0,1]$ and by $\tilde{A}[\alpha] = \left\{ x : \mu_{\tilde{A}}(x) \geq \alpha \right\}$ the α-level set of $\tilde{A}$ for $\alpha \in (0,1]$. Then $\tilde{A}[0]$ is the closure of the set $\left\{ x : \mu_{\tilde{A}}(x) > 0 \right\}$.

A fuzzy number $\tilde{a}$ is a fuzzy subset of R for which $\mu_{\tilde{A}}$ is a normal, upper-semicontinuous, fuzzy convex function with a compact support. Then for each $\alpha \in [0,1]$, the α-level set $\tilde{a}[\alpha]$ is a closed interval of the form $\tilde{a}[\alpha] = [\tilde{a}_L[\alpha], \tilde{a}_U[\alpha]]$, where $\tilde{a}_L[\alpha], \tilde{a}_U[\alpha] \in R$ and $\tilde{a}_L[\alpha] \leq \tilde{a}_U[\alpha]$.

If $+, -, \cdot, /$ is an operator for the fuzzy numbers (related to the equivalent operator $+, -, \cdot, /$ in the crisp case), then for the fuzzy numbers $\tilde{a}, \tilde{b}$, their

outcome is also a fuzzy number. Using α-cuts and the interval arithmetic we have:

$$(\tilde{a} + \tilde{b})[\alpha] = \left[\tilde{a}_L[\alpha] + \tilde{b}_L[\alpha], \tilde{a}_U[\alpha] + \tilde{b}_U[\alpha] \right]$$

$$(\tilde{a} - \tilde{b})[\alpha] = \left[\tilde{a}_L[\alpha] - \tilde{b}_U[\alpha], \tilde{a}_U[\alpha] - \tilde{b}_L[\alpha] \right]$$

$$(\tilde{a} \cdot \tilde{b})[\alpha] = [\min\{\tilde{a}_L[\alpha]\tilde{b}_L[\alpha], \tilde{a}_L[\alpha]\tilde{b}_U[\alpha], \tilde{a}_U[\alpha]\tilde{b}_L[\alpha], \tilde{a}_U[\alpha]\tilde{b}_U[\alpha]\},$$
$$\max\{\tilde{a}_L[\alpha]\tilde{b}_L[\alpha], \tilde{a}_L[\alpha]\tilde{b}_U[\alpha], \tilde{a}_U[\alpha]\tilde{b}_L[\alpha], \tilde{a}_U[\alpha]\tilde{b}_U[\alpha]\}],$$

$$(\tilde{a} / \tilde{b})[\alpha] = [\min\{\tilde{a}_L[\alpha]/\tilde{b}_L[\alpha], \tilde{a}_L[\alpha]/\tilde{b}_U[\alpha], \tilde{a}_U[\alpha]/\tilde{b}_L[\alpha], \tilde{a}_U[\alpha]/\tilde{b}_U[\alpha]\},$$
$$\max\{\tilde{a}_L[\alpha]/\tilde{b}_L[\alpha], \tilde{a}_L[\alpha]/\tilde{b}_U[\alpha], \tilde{a}_U[\alpha]/\tilde{b}_L[\alpha], \tilde{a}_U[\alpha]/\tilde{b}_U[\alpha]\}],$$

if $\tilde{b}[\alpha]$ does not contain zero for all $\alpha \in [0,1]$ in the last case.

A fuzzy number $\tilde{a}$ is called positive ($\tilde{a} \geq 0$) if $\mu_{\tilde{a}}(x) = 0$ for x<0 and it is called strictly positive ($\tilde{a} > 0$) if $\mu_{\tilde{a}}(x) = 0$ for x≤0.

An L-R fuzzy number is a fuzzy number with the membership function of the form:

$$\mu_{\tilde{a}}(x) = \begin{cases} L\left(\dfrac{x-a}{b-a}\right), & x \in [a,b] \\ 1, & x \in [b,c] \\ R\left(\dfrac{d-x}{d-c}\right), & x \in [c,d] \\ 0, & otherwise \end{cases},$$

where $L, R : [0,1] \rightarrow [0,1]$ are non-decreasing functions such that $L(0) = R(0) = 0$ and $L(1) = R(1) = 1$.

A triangular fuzzy number, denoted further by $[a,b,c]$, is a L-R number with the membership function of the form:

$$\mu_{\tilde{a}}(x) = \begin{cases} \dfrac{x-a}{b-a}, & x \in [a,b] \\ \dfrac{x-c}{b-a}, & x \in [b,c] \\ 0, & otherwise \end{cases}.$$

In the following, to approximate a desired fuzzy output $\tilde{f}$, we check monotonicity of the underlying function $f(x)$, when all of the parameters of the whole model, except of the single argument x, are held fixed.

The same idea, related to the extension principle, applies also for $\tilde{\lambda} > 0$ given by a L-R fuzzy number, where $\tilde{\lambda}$ is the fuzzy intensity parameter of the exponential probability density function. Therefore, if $f(\lambda)$ is an increasing function, then for the given α, the left end point $\tilde{f}_L[\alpha]$ is approximated using the crisp value $\tilde{\lambda}_L[\alpha]$ as the intensity of the exponential random variables generated in Monte Carlo simulations. In the same way $\tilde{\lambda}_U[\alpha]$ is used to approximate $\tilde{f}_U[\alpha]$ (see also, e.g., [2, 23, 24]).

4.2. Simulations in the fuzzified environment

We assume that the parameters of each type of the connection, namely $R^{(k)}$, $\Lambda^{(k)}$, $c_{const,rpl}^{(k)}$, $c_{const,rpr}^{(k)}$, $c_{var,rpl}^{(k)}$, $c_{var,rpr}^{(k)}$ for $k=1,\dots,K$, are given by L-R numbers. Therefore in the following they are denoted by $\tilde{R}^{(k)}$, $\tilde{\Lambda}^{(k)}$ etc. Then for the intensity matrix of the transitions we have $\tilde{\Lambda}^{(k)} = \left(\tilde{\lambda}_{kl}^{(j)}\right)_{k,l \in S}$, i.e. such a matrix consists of

five rows and five columns of positive fuzzy numbers and the pattern of strictly positive entries is the same as for (6).

Using the Monte Carlo simulations, the output similar to the one discussed in Section 3.2 is then obtained. But now these values, like the discounted value of the maintenance costs, are given by fuzzy numbers, so notation like $\widetilde{PV}(c)$ will be used further.

The fuzzy output is approximated by α-level sets $\widetilde{PV}(c)[\alpha]$ evaluated using Algorithm 2 (see Appendix). Firstly, we should decide if the left or right end points of $\widetilde{PV}(c)$ should be estimated during the simulations. Then a binary variable *left*=1 or *left*=0 is set, respectively. The other parameters are: a starting value $\alpha \in (0,1]$, an upper bound $\alpha_1 \in (\alpha_0,1]$ and an increment $\Delta\alpha>0$.

In the first phase, for the given α, the relevant α-level cuts of the fuzzy parameters for all types of connections are found. Based on the monotonicity of $PV(c)$, for each possible fuzzified argument and the fixed value of *left*, the left or right end point of the α-level cut of each parameter is then selected.

During the second step, the simulations of the whole WDS are conducted using Algorithm 1 (see Appendix). The generated crisp output $PV(c)$ is approximation of $PV_L(c)[\alpha]$ or $PV_U(c)[\alpha]$ depending on the value of *left*.

To obtain the mentioned approximation of the whole fuzzy output, the value of α in the above procedure is gradually set to subsequent values, starting from α_0 up to α_1 with the given increment $\Delta\alpha$. After evaluation of the left end points of $PV(c)$ (for *left*=1), the right end points are found in the same manner (i.e. *left*=0 is applied). Then the obtained intervals are putting on one another to form the approximated fuzzy number.

The above method is based on the extension principle introduced by Zadeh (see [31]) and a similar approach using α-level cuts is also applied in financial mathematics for pricing of the derivatives (see, e.g., [23, 24]), in optimization of queuing systems (see, e.g., [2]) etc.

Because the considered function for the fuzzy result (e.g. $\widetilde{PV}(c)$ is not given by an exact, analytical formulae, the simulations are required for its evaluation. As previously mentioned, it should be determined for each of the fuzzy parameter, if the left or the right end point of its α-level cut is used to calculate the left or right end point of the output interval during the first step of the above procedure. This choice is strictly related to the monotonicity of a function $f(p)$ of the considered outcome, i.e. if the output (like the discounted maintenance costs) is an increasing or decreasing function of the given parameter p (like the constant costs of repair). In Algorithm 2 (see Appendix) the considered approach is applied for $\widetilde{PV}(c)$, but the same method can be used for other output functions $f(.)$, which are important from the practical point of view.

4.2.1. Example III

For simplicity, in the following, we assume that the fuzzy parameters of the considered single type of the connection are given by some triangular numbers. However, other types of L-R numbers can be also directly used in our approach. Let us start from the set of the parameters:

$$\tilde{c}_{const,rpl}^{(1)} = [4,5,6], \tilde{c}_{const,rpr}^{(1)} = [3,4,5], \tilde{c}_{var,rpl}^{(1)} = [2,3,4], \tilde{c}_{var,rpr}^{(1)} = [1,2,3].$$
$$(15)$$

Other parameters are the same as in Section 3.2.1. It is easily seen from (15), that only the constant and variable costs of repairs and replacements are strictly fuzzy triangular numbers in the considered case. Therefore, they are the only source of uncertainty in the following example.

Using the Monte Carlo approach described by Algorithm 2, the fuzzy discounted maintenance costs $\widetilde{PV}(c)$ for the Vasicek model (see Figure 4, circles) and the fuzzy average of the nominal costs of replacements $\tilde{c}_{rpl}$ (see Figure 5) are approximated. The obtained fuzzy values are almost triangular symmetrical fuzzy numbers for which the 1-level cut sets are equal to the output estimated in Section 3.2.2. Some approximated values of the intervals for different α can be found in Table 2.

Table 2. Exemplary α-level cuts of the fuzzy discounted maintenance costs $\widetilde{PV}(c)$ and the fuzzy average of the nominal costs of replacements $\tilde{c}_{rpl}$ from Example III

α	$\widetilde{PV}(c)$	$\tilde{c}_{rpl}$
0	[1208.23, 1949.52]	[4.16594, 6.33188]
0.1	[1245.29, 1912.46]	[4.27424, 6.22358]
0.5	[1393.55, 1764.2]	[4.70742, 5.79039]
0.9	[1541.81, 1615.94]	[5.14061, 5.3572]

Both $PV(c)$ and c_{rpl} are increasing functions of the fuzzy parameters (15). Therefore, to calculate the left end point of the interval $PV_L(c)[\alpha]$ (or the right end point $PV_U(c)[\alpha]$, respectively) for the given value of α, only the left end points (or the right points) of the α-level cuts for (15) should be considered. The same applies for $\tilde{c}_{rpl}$.

Using the Monte Carlo simulations, the present values of the maintenance costs can be also found for the more classical models of the interest rates, i.e. the nominal costs (see Figure 4, squares) and the constant yield (see Figure 4, rhombus). The obtained approximations of the fuzzy numbers are similar in shape, but significantly different.
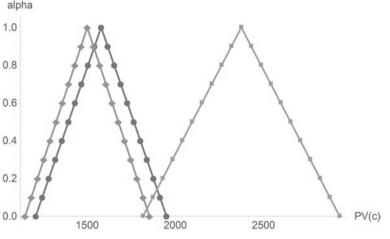

Fig. 4. Fuzzy discounted and nominal maintenance costs from Example III

4.2.2. Example IV

Of course, not only the maintenance costs can be considered as uncertain values in the practical situations. Therefore we analyse situation if the unconditional replacement age is also given by a triangular, fuzzy number, denoted further by $\tilde{R}^{(1)}$. Other fuzzy parameters (namely $\tilde{c}_{const,rpl}^{(1)}$, $\tilde{c}_{const,rpr}^{(1)}$, $\tilde{c}_{var,rpl}^{(1)}$, $\tilde{c}_{var,rpr}^{(1)}$) are described by (15) and the crisp parameters are the same as in Section 3.2.1.

As previously, to approximate a desired fuzzy outcome, we should check the monotonicity of the considered function for the introduced fuzzy parameters. As noted in Section 4.2.1, $PV(c)$ is an increasing
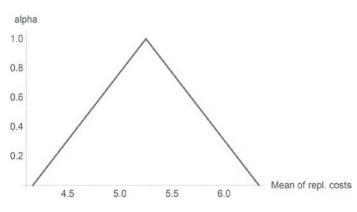
Fig. 5. Fuzzy average (nominal) costs of replacements from Example III

function of the costs, but the dependency on $R^{(1)}$ is not so straightforward. It is possible, that for the higher values of $R^{(1)}$, the overall maintenance costs will increase. However, for the considered values of the parameters, as emphasized by Figure 1, such a relation is a decreasing function. Therefore, to evaluate $PV_L(c)[\alpha]$, the left end points of the α-level cuts of the fuzzy costs and the right end point of the α-level cut of $\tilde{R}^{(1)}$ (given by $R_U^{(1)}([\alpha])$) should be taken into account.

The obtained approximations of $\widetilde{PV}(c)$ are shown in Figure 6, where x axis is the present value of the maintenance costs, y axis is related to the triangular values $\tilde{R}^{(1)}$ of the form $[y-2, y, y+2]$ and z
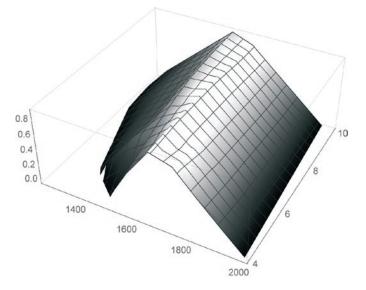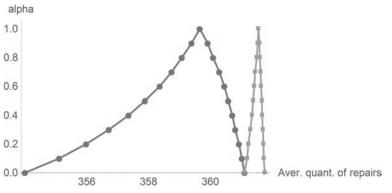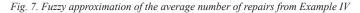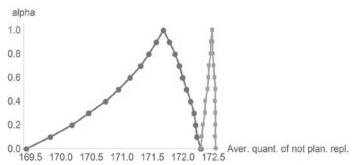

Fig. 6. Fuzzy discounted maintenance costs from Example IV

Fig. 7. Fuzzy approximation of the average number of repairs from Example IV

Fig. 8. Fuzzy approximation of the average number of not planned replacements from Example IV
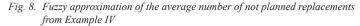
axis is the relevant α-level cut. As it can be seen, the estimated fuzzy values are L-R fuzzy numbers for which the supports are wider, if y is higher in the formulae $\tilde{R}^{(1)} = [y-2, y, y+2]$.

As in Example II, the fuzzy characterizations, which are important for the practitioners, i.e. the average number of repairs $\tilde{x}_{rpr}$ (see Figure 7) and the average number of not planned replacements $\bar{x}_{rpl,NR}$ (see Figure 8) are also obtained. Then, the relevant outputs for $\tilde{R}^{(1)} = [2,4,6]$ (Figures 7 and 8, circles) and $\tilde{R}^{(1)} = [6,8,10]$ (Figures 7 and 8, squares) can be compared. As it can be seen, for the higher value of $\tilde{R}^{(1)}$, the obtained L-R number is significantly shifted right with a shorter support.

4.2.3. Example V

All of the parameters of the connections in the WDS (i.e. the first group mentioned in Section 2.4) can be fuzzified. Such an approach reflects various sources of uncertainties which results in the necessary incorporation of the experts' knowledge. Then, in the following example, apart from applying the fuzzy values:

$$\tilde{R}^{(1)} = [8,10,12], \tilde{c}_{const,rpl}^{(1)} = [4,5,6], \tilde{c}_{const,rpr}^{(1)} = [3,4,5], \tilde{c}_{var,rpl}^{(1)} = [2,3,4], \tilde{c}_{var,rpr}^{(1)} = [1,2,3]$$

(16)

also the intensity matrix of transitions $\tilde{\Lambda}^{(k)}$ is described by triangular fuzzy numbers, so that:

$$\tilde{\Lambda}^{(1)} = \begin{pmatrix} 0 & 0 & [10,12,14] & 0 & 0 \\ 0 & 0 & [22,24,26] & [24,26,28] & 0 \\ [0.4,0.5,0.6] & [0.9,1,1.1] & 0 & [0.9,1,1.1] & 0 \\ [0.3,0.4,0.5] & [0.8,0.9,1] & 0 & 0 & [0.2,0.3,0.4] \\ [0.5,0.6,0.7] & [1,1.1,1.2] & 0 & 0 & 0 \end{pmatrix}.$$

(17)

The fuzzy values, given by (17), describe only the fuzzy intensities of the exponential distributions of the transitions between the states from the set S, as discussed in Section 4.1. It is easily seen, that the fuzzy numbers (17) are "close" to the crisp values (13) assumed in Example I. Therefore, the fuzzy output obtained during current analysis can be easily compared with the values from the previously considered examples.

As previously mentioned, in order to evaluate the fuzzy output, it is necessary to check the monotonicity of the function for the considered parameter. And $PV(c)$ is a decreasing function of $\lambda_{02}^{(1)}$, $\lambda_{12}^{(1)}$, $\lambda_{13}^{(1)}$, because for the higher values of these intensities, the expected time of replacement or repair is lower, so the final cost is also lower. All of the other parameters can be examined in the same way.
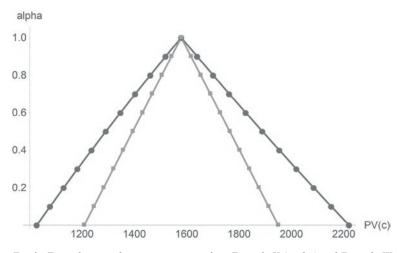
Fig. 9. Fuzzy discounted maintenance costs from Example V (circles) and Example III (squares)
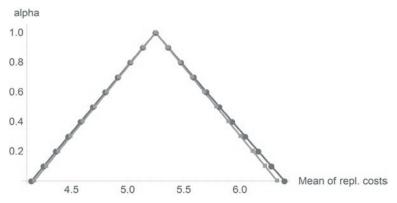

Fig. 10. Fuzzy average (nominal) costs of replacements from Example V (circles) and Example III (squares)
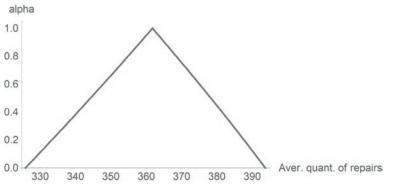

Fig. 11. Fuzzy approximation of the average number of repairs from Example V

Fig. 12. Fuzzy approximation of the average number of not planned replacements from Example V
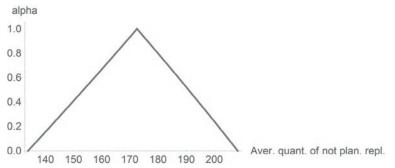
Table 3. Exemplary α-level cuts of the fuzzy discounted maintenance costs $\widetilde{PV}(c)$ and the fuzzy average of the nominal costs of replacements $\tilde{c}_{rpl}$ from Example V

α	$\widetilde{PV}(c)$	$\tilde{c}_{rpl}$
0	[1026.98, 2220.02]	[4.14232, 6.39796]
0.1	[1077.91, 2152.23]	[4.25161, 6.28049]
0.5	[1291.05, 1889.06]	[4.6915, 5.81668]
0.9	[1519.32, 1639.42]	[5.13678, 5.36153]

The approximated fuzzy discounted maintenance costs, which are evaluated using the Monte Carlo simulations are also compared with the similar output obtained in Example III (see Figure 9). As it is seen, in the considered case $\widetilde{PV}(c)$ is a L-R fuzzy number, almost a triangular one, with a wider support comparing to the fuzzy number from Example III.

In a similar way, the fuzzy average (nominal) costs of replacements is approximated. In Figure 10, the obtained values are compared with the similar fuzzy output from Example III. Some of the evaluated intervals for different $α$ can be found in Table 3.

As previously, the fuzzy approximations of the average number of repairs $\tilde{x}_{rpr}$ (see Figure 11) and the average number of not planned replacements $\tilde{x}_{rpl,NR}$ (see Figure 12) are also obtained using the introduced approach. These fuzzy L-R numbers are almost triangular, but their supports are wider comparing to the outputs from the examples, where the intensities are strictly crisp values.

5. Conclusions

In this paper the model of the transitions between the states of the single connection of the WDS is proposed. This model is related to a semi-Markov process and the deterministic jumps caused by the unconditional replacement age for each pipeline. Using the Monte Carlo approach, the behaviour of the whole WDS is simulated and the costs of the maintenance services (i.e. repairs and replacements) are evaluated. Then, based on a stochastic process (the one-factor Vasicek model of interest rates), the present value of the future cash flows and other important statistical measures of the mentioned costs are calculated. The estimated differences between the classical approach (i.e., a constant yield) and the proposed model are emphasized. Apart from the strictly crisp setup, the fuzzification of some parameters of the connection is considered. An introduction of the fuzzy numbers leads to a better incorporation of the experts' knowledge and more real-life modelling of the uncertain parameters. The necessary numerical algorithms and some relevant examples of the simulated output for both the crisp and the fuzzy settings are also provided.

There are various possible fields for future works based on the approach introduced in this paper. Firstly, other kinds of distributions, instead of the exponential ones, can be assumed – e.g., the repairing time can be described by some lognormal random variable. Secondly, if a different cdf is applied, a relevant case can be calibrated using some real-life data, so statistical tests concerning a validity of this cdf (versus, e.g., the exponential one) can be proposed. It also leads to a possibility of measuring a level of a difference in the estimated maintenance costs, which can be important for the practitioners.

Appendix

Algorithm 1: Estimation of the maintenance costs – the crisp case.

Input: *The number of simulations n, the parameters of each type k of the connection, the parameters of the interest rate model (see Section 2.4).*

Output: *The present value of the maintenance costs PV(c).*

for *i*=1 **to** *n* **do**

    for *k*=1 **to** *K* **do**

        for *j*=1 **to** n_k **do**

            Generate an independent trajectory $S^{(j)}$ of the process of the states (see Section 2.1);

            Store the times t_l of transitions to the states 0 or 1 and the relevant periods of the repairs and the replacements (see Section 2.2);

            Evaluate and store all of the costs $c^j(t_l)$ for each t_l (see Section 2.2);

        end

    end

    Put all of the times $t_1, t_2, \ldots$ in the increasing order $t_{(1)} \leq t_{(2)} \leq \ldots \leq t_{(m)} \leq T$;

    PV_i=0;

for *j*=1 **to** *m* **do**

    Generate $r_{t_{(j)}}$ and evaluate the discounted costs $PV\left(c_{t_{(j)}}\right)$ (see Section 2.3);

    $PV_i = PV_i + PV\left(c_{t_{(j)}}\right);$

    end

end

$$PV(c) = \frac{1}{n}\sum_{i=1}^{n} PV_i \, ;$$

Algorithm 2: Estimation of the maintenance costs – the fuzzy case.

Input: The number of simulations *n*, the fuzzy parameters of each type *k* of the connection (see Section 4.2), the crisp parameters of the interest rate model (see Section 2.4), the starting value α_0, the upper bound α_1, the increment $\Delta\alpha > 0$, the binary variable *left*.

Output: The left or right end points of approximation of the fuzzified present value of the maintenance costs $\widetilde{PV}(c)$.

$\alpha = \alpha_0;$

while $\alpha < \alpha_1$ **do**

    for *k*=1 **to** *K* **do**

For the k type of the connection, find α-level cuts of the fuzzy parameters: $\tilde{R}^{(k)}[\alpha] = \left[\tilde{R}_L^{(k)}[\alpha], \tilde{R}_U^{(k)}[\alpha] \right]$, $\tilde{\Lambda}^{(k)}[\alpha] = \left[\tilde{\Lambda}_L^{(k)}[\alpha], \tilde{\Lambda}_U^{(k)}[\alpha] \right]$ *etc.;*

for *each fuzzified parameter p* **do**

    Check monotonicity of PV(c) for the argument p;

    Set $p_L[\alpha]$ or $p_U[\alpha]$ depending on the monotonicity of PV(c) for the operator left;

  end

end

Use Algorithm 1 to generate the crisp value of PV(c);

if *left* **then**

    $PV_L[\alpha]=PV(c)$;

else

    $PV_U[\alpha]=PV(c)$;

end

  $\alpha = \alpha + \Delta\alpha$;

end

References

1. Amani N, Ali N M, Mohammed A H, Samat R A. Maintenance and management of wastewater system components using the condition index system, prediction process and costs estimation. Eksploatacja i Niezawodnosc – Maintenance and Reliability. 2013; 15(2): 161–168.
2. Buckley J J. Simulating Fuzzy Systems. Berlin Heidelberg: Springer-Verlag, 2005, http://dx.doi.org/10.1007/b100371.
3. Clark R M, Thurnau R C. Evaluating the risk of water distribution system failure: A shared frailty model. Front. Earth Sci. 2011; 5(4): 400–405.
4. Fadaee M J, Tabatabaei R. Estimation of failure probability in water pipes network using statistical model. Engineering Failure Analysis. 2011; 18: 1184–1192, http://dx.doi.org/10.1016/j.engfailanal.2011.02.013.
5. Glasserman P. Monte Carlo Methods in Financial Engineering. New York: Springer, 2004.
6. Gonzalez A, Pons O, Vila M A. Dealing with uncertainty and imprecision by means of fuzzy numbers. International Journal of Approximate Reasoning. 1999; 21: 233–256, http://dx.doi.org/10.1016/S0888-613X(99)00024-9.
7. Goulter I C, Davidson J, Jacobs P. Predicting watermain breakage rates. J. Water Resour. Plann. Manage. ASCE. 1993; 119(4): 419–436, http://dx.doi.org/10.1061/(ASCE)0733-9496(1993)119:4(419).
8. Hryniewicz O. Comparison of Fuzzy and Crisp Random Variables by Monte Carlo simulations. In: Grzegorzewski P, Gągolewski M, Hryniewicz O, Gil M A. (eds.). Strengthening Links Between Data Analysis and Soft Computing. Springer, 2015, http://dx.doi.org/10.1007/978-3-319-10765-3_2.
9. Hryniewicz O, Kaczmarek K, Nowak P. Bayes statistical decisions with random fuzzy data – An application for the Weibull distribution. Eksploatacja i Niezawodnosc – Maintenance and Reliability. 2015; 17(4): 610–616, http://dx.doi.org/10.17531/ein.2015.4.18.
10. Islam M S, Sadiq R, Rodriguez M J, Najjaran H, Francisque A, Hoofar M. Water distribution system failure: a framework for forensic analysis. Environ. Syst. Decis. 2014; 34: 168–179, http://dx.doi.org/10.1007/s10669-013-9464-3.
11. Kanakoudis V K, Tolikas D K. Assessing the performance level of a water system. Water, Air, and Soil Pollution. 2004; 4: 307–318, http://dx.doi.org/10.1023/B:WAFO.0000044807.41719.c7.
12. Kleiner Y, Adams B J, Rogers J S. Long-term planning methodology for water distribution system rehabilitation. Water Resources Research. August 1998; 34(8): 2039–2051, http://dx.doi.org/10.1029/98WR00377.
13. Kumar G, Jain V, Gandhi O P. Steady-state availability analysis of repairable mechanical systems with opportunistic maintenance by using semi-Markov process. Int J Syst Assur Eng Manag. 2014; 5(4): 664–678, http://dx.doi.org/10.1007/s13198-014-0231-8.
14. Lee K H. First Course on Fuzzy Theory and Applications. Berlin Heidelberg: Springer, 2005.
15. Lindhe A, Rosén L, Norberg T, Bergstedt O. Fault tree analysis for integrated and probabilistic risk analysis of drinking water systems. Water Research. 2009; 43: 1642–1653, http://dx.doi.org/10.1016/j.watres.2008.12.034.
16. Liu Z, Kleiner Y, Rajani B, Wang L, Condit W. Condition Assessment Technologies for Water Transmission and Distribution Systems. Washington DC: U.S. Environmental Protection Agency, 2012; EPA/600/R-12/017.
17. Mailhot A, Poulin A, Villeneuve J P. Optimal replacement of water pipes. Water Resour. Res. 2003; 39(5): HWC 2-1 to HWC 2-14.
18. Malinowski J. Finding reliability measures for a water supply network by means of component state factorization and series-parallel aggregation. http://www.ibspan.waw.pl/woda-dla-slaska/publikacje.htm

19. Malinowski J. Reliability oriented simulation of a water supply system, with the use of a components database. In: Wittmann J, Wieland R. (eds.) Simulation in Umwelt- und Geowissenschaften. Workshop Müncheberg 2015. Aachen: Shaker Verlag, 2015.

20. Marchionni V, Cabrala M, Amadoa C, Covasa D. Water supply infrastructure cost modelling. Procedia Engineering. 2015; 119: 168–173, http://dx.doi.org/10.1016/j.proeng.2015.08.868.

21. Neelakantan T R, Suribabu C R, Lingireddy S. Optimisation procedure for pipe-sizing with break-repair and replacement economics. Water SA. April 2008; 34(2): 217–224.

22. Norris J R. Continuous-time Markov chains. Cambridge University Press, 1997, http://dx.doi.org/10.1017/CBO9780511810633.

23. Nowak P, Romaniuk M. A fuzzy approach to option pricing in a Levy process setting. International Journal of Applied Mathematics and Computer Science. 2013; 23(3): 613–622, http://dx.doi.org/10.2478/amcs-2013-0046.

24. Nowak P, Romaniuk M. Application of Levy processes and Esscher transformed martingale measures for option pricing in fuzzy framework. Journal of Computational and Applied Mathematics. 2014; 263: 129–151, http://dx.doi.org/10.1016/j.cam.2013.11.031.

25. Rokstad M M, Ugarelli R M. Minimising the total cost of renewal and risk of water infrastructure assets by grouping renewal interventions. Reliability Engineering and System Safety. 2015; 142: 148–160, http://dx.doi.org/10.1016/j.ress.2015.05.014.

26. Romaniuk M. Application of Markov chain and interest rate process for forecasting of costs of maintenance of pipelines. In: Wittmann J, Wieland R. (eds.) Simulation in Umwelt- und Geowissenschaften. Workshop Müncheberg 2015. Aachen: Shaker Verlag, 2015.

27. Romaniuk M, Nowak P. Monte Carlo methods: theory, algorithms and applications to selected financial problems. Warszawa: ICS PAS, 2015.

28. Scheidegger A, Leitão J P, Scholten L. Statistical failure models for water distribution pipes – A review from a unified perspective. Water Research. 2015; 83: 237–247, http://dx.doi.org/10.1016/j.watres.2015.06.027.

29. Shamir U, Howard C D D. Analytical approach to scheduling pipe replacement. J. Am. Water Works Assoc. 1979; 74(3): 140–147.

30. Williams J, Steele N, Robinson H. Reasoning with Non-Numeric Linguistic Variables. Journal of Computing and Information Technology. 2002; 10(4): 261–281, http://dx.doi.org/10.2498/cit.2002.04.02.

31. Zadeh L A. Fuzzy sets. Information and Control. June 1965; 8(3): 338–353, http://dx.doi.org/10.1016/S0019-9958(65)90241-X.

32. Zimoch I, Szymik-Gralewska J. Risk assessment methods of a water supply systems in terms of reliability and operation costs. In: Membretti S, Brebbia C A. (eds.) Urban Water II. Southampton: WIT Press, 2014, http://dx.doi.org/10.2495/UW140051.

Maciej ROMANIUK
Systems Research Institute Polish Academy of Sciences
ul. Newelska 6, 01-447 Warszawa, Poland
and
The John Paul II Catholic University of Lublin,
ul. Konstantynów 1 H, 20-708 Lublin, Poland
e-mail: mroman@ibspan.waw.pl