Statistical failure models for water distribution pipes – a review from a unified perspective

Andreas Scheidegger[a,*], João P. Leitão[a], Lisa Scholten[a]

[a] *Eawag Swiss Federal Institute of Aquatic Science and Technology Uberlandstrasse 133 8600 Dübendorf Switzerland*

Abstract

This review describes and compares statistical failure models for water distribution pipes in a systematic way and from a unified perspective. The way the comparison is structured provides the information needed by scientists and practitioners to choose a suitable failure model for their specific needs.

The models are presented in a novel framework consisting of: 1) Clarification of model assumptions. The models originally formulated in different mathematical forms are all presented as failure rate. This enables to see differences and similarities across the models. Furthermore, we present a new conceptual failure rate that an optimal model would represent and to which the failure rate of each model can be compared. 2) Specification of the detailed data assumptions required for unbiased model calibration covering the structure and completeness of the data. 3) Presentation of the different types of probabilistic predictions available for each model.

Nine different models and their variations or further developments are presented in this review. For every model an overview of its applications published in scientific journals and the available software implementations is provided.

The unified view provides guidance to model selection. Furthermore, the model comparison presented herein enables to identify areas where further research is needed.

Keywords: Pipe failure model, failure rate, break model, deterioration model

*Corresponding author andreas scheidegger@eawag ch

1. Introduction

1 1 Need for deterioration models

The structural condition of urban water distribution infrastructures is important for the continuity and quality of the water distribution services provided by these systems. The financial investments needed for the rehabilitation, adaptation, and expansion of existing urban water systems (incl. water treatment as well as drainage and wastewater treatment) are estimated at 1% of the annual GDP in the OECD member states, rehabilitation accounting for up to half of the total needs (OECD, 2006).

Targeted research programs in e.g. Canada[1], Australia[2], the United States[3], and Europe[4] also acknowledged the need for approaches to assess the deterioration and failure development of urban water distribution networks. This is because pipe deterioration may have a significant impact on some of the fundamental objectives of water distribution networks, e.g. reliability and continuity of service. It is important to be able to predict future deterioration in order to determine the optimal amount and timing of the required rehabilitation efforts. These are central inputs for technical asset management and defining the long-term budgets. Knowledge about how the structural condition of pipes develops over time is key to designing and choosing replacement and maintenance strategies.

A range of software has been proposed to support water supply infrastructure asset management based on pipe deterioration models (Burn et al., 2003; Cardoso et al., 2012; Kleiner and Rajani, 2010; Le Gat, 2009; Renaud et al., 2012; Saegrov, 2005). This allows, for example, comparing rehabilitation strategies based on key performance indicators, such as “main failures” or “water resources availability” (Alegre et al., 2006). Any rehabilitation strategy can be defined by, for example, incorporating constraints such as budget or

[1] Several projects of the National Research Council Canada (http://www.nrc-cnrc.gc.ca/eng/achievements/highlights/2008/aging_water_systems.html)

[2] Sustainable Asset Management Program developed by CSIRO (http://www.webcitation.org/6UB8dXfP3)

[3] Aging Water Infrastructure Program of the US Environmental Protection Agency (http://www.webcitation.org/6UB8jPf2P) several projects of the Water Research Foundation Asset Management (http://www.webcitation.org/6UB8nn7tE)

[4] European Union FP5 project CARE-W (http://www.webcitation.org/6UB8sNy2c) several follow-up projects by different funding agencies AWARE-P (http://www.webcitation.org/6UB8xroUh) EU FP7 project i-TRUST (http://www.webcitation.org/6UB9662Zn)

work load, while the deterioration models should represent the deterioration behavior of the pipes as realistically as possible.

Even though modeling the deterioration of water distribution and drainage systems shares some challenges such as frequently incomplete data sets, its modeling is different due to the specific characteristics of each system and to the unique characteristics of the available data. Water distribution system data contain information about when events (e.g. pipe failures) occurred while drainage system data usually provide information about the condition of the pipes at the time of the inspection. Therefore we do not include statistical sewer deterioration models into this review. In this review, only statistical pipe failure models (as defined in the next subsection) developed for water distribution systems are considered.

1 2 Aim of this review and models covered

The aim of this review is to describe and compare statistical pipe failure models in a systematic way. This is intended to support practitioners and scientists in choosing a suitable statistical pipe failure model to support pipe rehabilitation and asset management decisions as well as to identify research needs. Available reviews by Kleiner and Rajani (2001), Liu et al. (2012), and to a very limited extent also Nishiyama and Filion (2013) and St. Clair and Sinha (2012) (adding information regarding artificial neural networks and fuzzy logic models) give a broad overview of statistical pipe failure models. These reviews, however, do not provide the information needed for objective model characterization, comparison and selection as recognized by Liu et al. (2012).

Instead, we discuss statistical pipe failure models from a novel unified perspective consisting of: i) a clarification of the model assumptions independent of how the models are expressed mathematically in the original publications (Section 3.1), ii) a specification of *detailed* data assumptions for model calibration covering the structure and completeness of the data (Section 3.2), and iii) a presentation of the type of probabilistic predictions published (Section 3.3). We further provide references to illustrative applications and available software implementations as published in the literature. For the first point we mathematically reformulated the models to represent them by their failure rates. We present a novel conceptual failure rate in Section 3.1 that includes all desired properties to which the failure rates of the models can be compared to. In the second point the often only implicitly assumed data characteristics important for model calibration are discussed. The third point deals with the presented predictions for each model as obtaining pipe failure predictions is usually the major motivation for using pipe failure models and the formulation of appropriate predictive distributions requires care.

This paper is organised as follows: Section 2 describes the different types of pipe failure models; this section is important to underline why we have focused the review on the *statistical pipe failure* models. A thorough explanation of model properties is presented in Section 3, allowing the structure of the model review presented in Section 4 to be understood. The core of this review are Section 4 and Table 1; the models are critically discussed, focusing on the following model properties: structure, calibration, predictions, further developments, applications and software implementations. Section 5 contains a discussion on how the review may assist model selection and potential for future developments is identified.

Readers interested only in the model properties and assumptions may want to proceed directly with Section 4 and consult Table 1. Readers looking for details, the structure of this review and the reasons for the proposed unified perspective may also read the introductory sections of this paper (Sections 2 and 3).

2. Types of deterioration models

Deterioration models for water distribution systems can be differentiated by at least three dimensions: the smallest described entity, the type of events that are modeled, and the modeled process. The smallest entity dimension relates to whether individual pipes (pipe models) or a pipe network (network models) are described. A model can describe different events: the occurrence of failures (failure models) or the end of the life span (lifetime or lifespan models). Finally, models aim to represent different processes, either by mimicking physical processes (physically-based models) or by attempting to describe the data generating process (statistical models). Any combination of these dimensions is possible, although from the available literature, only a part of these combinations is covered.

This three-dimension categorization (by entity, type of event, and process modeled) differs from other reviews, which are mostly based on a model categorization originating from Kleiner and Rajani (2001) (cf. section 2.3). The new categorization aims to avoid the following drawbacks of the Kleiner and Rajani (2001) characterization: i) it is based on the mathematical formulation of a model rather than on its underlying assumptions, ii) extensions with covariates are relatively simple so that "single-variate" models can always be extended to "multi-variate models" and are not a model limitation per se as discussed in Section 3.1, and iii) the "deterministic" category refers to statistical regression models which, strictly speaking, are also "probabilistic".

2 1 Individual pipe vs network models

Many of the earlier models describe the failure rate of a complete network (see Table 1 in Kleiner and Rajani, 2001). The network failure rate summarizes the overall condition of a network and is therefore an important performance indicator to assess the structural and operational condition (such as continuity of service) of the whole system. Predicting directly how the failure rate of a complete network changes is difficult because technical management actions are generally performed at pipe level. Furthermore, the age and material distributions of the pipe network change over time so that a simple extrapolation of the failure rate is not adequate. To consider this, it is necessary to model the behavior of the pipes individually. The system wide failure rate can then be derived by aggregating the individual predictions. For this reason, we only consider pipe-based deterioration models in further detail within this review.

2 2 Failure vs lifetime (or lifespan) models

Both the failure behavior and the lifespan of pipes depend on deterioration processes. Models describing the lifespan are less flexible than failure models because the definition of a lifespan implies a combination of pipe deterioration and the management strategy that defines when a pipe reaches the end of its (e.g. economic, operational,...) life (Le Gat et al., 2013; Scholten et al., 2013). The problem with lifespan models is that they do not assess pipe deterioration, but rather when a pipe has been replaced in the past due to some reason. This may often be due to other reasons than structural condition, e.g. because of collaborative works with other infrastructure sectors or because of the available budget to be spent within a predefined time period. Therefore, the results of such models can neither be compared across networks nor are they useful to assess different future management strategies. Pipe failure models do not have these limitations: the failure behavior is modeled in isolation, which can then be combined with any management strategy. Assessing how the number of pipe failures in a network evolves under different management strategies is one of the most important and practically relevant questions in rehabilitation planning. As this requires pipe failure models, we only consider these in this review.

2.3. Statistical vs. physically-based models

Pipe deterioration models either aim to model the occurrence of the resulting failures statistically or describe the physical deterioration processes (such as corrosion) directly.

Statistical models and their historical development are extensively reviewed in Kleiner and Rajani (2001). This review has been recently updated in a report to US EPA (Liu et al., 2012). Kleiner and Rajani differentiate three model categories: *deterministic*, *probabilistic multi-variate* and *probabilistic single-variate*. This is based on whether models accommodate input and parameter uncertainties (if true: *probabilistic*, otherwise: *deterministic*) and whether several influencing covariates apart from pipe age are considered (if true: *multi-variate*, otherwise: *single-variate*). Liu et al. (2012) have added the "type of deterioration" to this differentiation, i.e. whether the condition is rated or a breakage frequency calculated. Artificial intelligence models, such as artificial neuronal networks, are also statistical models as they can be considered as highly flexible non-linear regression models (with the difference that most model assumptions cannot be expressed explicitly).

An extensive review of the physically-based models is given by Rajani and Kleiner (2001), again with an update in Liu et al. (2012). These models often require the (costly) acquisition of data on individual pipe and environmental characteristics. This is generally only feasible for a few highly critical, large-diameter transmission pipes, but not for the whole distribution network (Kleiner and Rajani, 2001). For this reason, we limit the review to statistical models.

3. Properties of statistical pipe failure models

3.1. Formulating statistical pipe failure models

Statistical pipe failure models describe the occurrence of failures as a stochastic process, as depicted in Figure 1. Mathematically it is not important what a failure event represents. Often failures are defined as events that require an immediate action, such as pipe bursts. However, other definitions are possible depending on the available data. Also, failures are not distinguished by cause, i.e. failures due to deterioration of the pipe are treated similar to failures due to external random effects. The mental model behind Figure 1 assumes that i) a pipe can fail anytime, ii) repair happens immediately after a failure, and iii) pipe age and number of failures are unlimited. The first assumption is evident; a pipe failure can occur whenever a pipe is in service. The

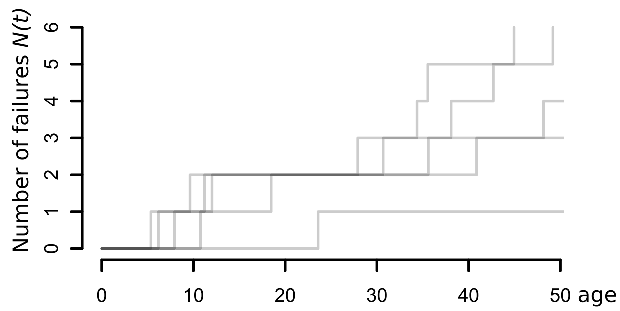


Figure 1 Several random realizations of the same stochastic failure process

second assumption is justified as typically the repair times are orders of magnitude shorter than the times between failures. Even though assumption three is not realistic, it is not a problem because the maximum pipe age and failure numbers are limited by the management model that triggers replacement.

Mathematically, there are several ways to describe such a stochastic failure process. The available pipe failure models describe it either by a failure rate, by the probability distributions of times between failures, or by the probability distribution of the age at occurrence of the ith failure (Le Gat, 2014). Depending on the assumptions, one representation is typically most convenient, even though any model can theoretically be expressed in all three ways (e.g. Cook and Lawless, 2010, chapter 2). For this reason, the mathematical representation is not regarded as a good foundation to classify or discuss the advantages and limitations of models, but rather a matter of mathematical convenience and personal preference.

Model structure

A system that may fail at any point in time can be characterized by a *failure rate* (also referred to as *hazard rate*, *intensity function*, or *rate of occurrence of failures*). The failure rate

$$\lambda(t \mid H(t)) = \lim_{\Delta \to 0^+} \frac{\text{Prob}\,(\text{failure in } [t, t+\Delta] \mid H(t))}{\Delta}$$

for age t is proportional to the probability that a failure occurs within an infinitesimally short time period from t conditional (depending) on the failure history $H(t)$. It is important to note that the failure rate is *not* a probability density, but rather a probability density per unit time.

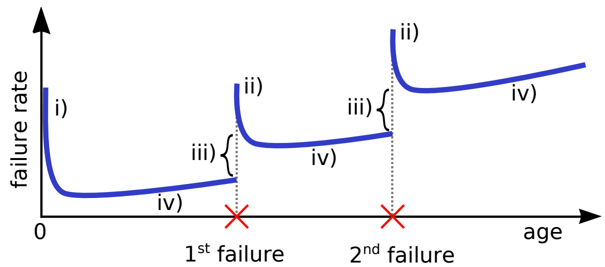


Figure 2 Conceptual failure rate of a drinking water pipe Four effects are illustrated i) increase due to installation failures ii) temporal increase after a failure because of the stress of the repair iii) persistent increase after each failure because the repair permanently weakened the structure iv) slow constant increase over time due to deterioration

For many technical systems, including drinking water pipes, "bath tub"-shaped failure rates are often assumed (Kleiner and Rajani, 2001). The rate is initially high due to stress and potential problems right after installation, e.g. caused by low-quality joint welding. It decreases quickly and is followed by a slow increase over time due to infrastructure deterioration. Since pipes are repairable, the failure rate is probably more complex. Figure 2 shows a conceptual failure rate for water distribution pipes with four expected characteristics: i) high failure rate at the beginning due to installation problems, ii) a sudden failure rate increase after every failure because of the stress of the repair, iii) a persistent increase after each failure because repairs weakened the structure permanently, and iv) a slow constant increase over time due to deterioration of the pipe. The spatial dependency of the failure rate on failures of other pipes in the system is not represented in Figure 2, although it has been reported (e.g. Kleiner and Rajani, 2001) that failures can trigger failures in neighboring pipes due to, for example, sudden pressure changes in the system.

A model that represents all these features would be vastly complex and the parameter estimation from failure data extremely challenging. Therefore, all published models include simplifications, assuming a failure rate that replicates only parts of the conceptual failure rate presented in Figure 2. General statements on how severely such simplifications influence the performance of the models cannot be made in the absence of a full representation of the conceptual model.

Every model—independent of its mathematical description—has a corresponding representation of the failure rate. We consider a comparison of the explicitly or implicitly assumed failure rate to be a good basis to discuss and visualize different model assumptions. The actual mathematical expression of the failure rate is not necessarily easy to interpret, especially for models originally not described by the failure rate. Therefore, a visual representation of the failure rate for each model considered in this review is presented in Table 1. The core assumptions, such as whether the failure rate depends on the number of previous failures or on the pipe age, are also stated.

Model covariates

In this review we do not focus on what covariates are explicitly considered in each model. This is mainly because any characteristics with sufficiently complete records can be introduced as model covariates (also called *covariables* or *explanatory variables*).

The selection of covariates is a complex task as it depends on the predictions of interest, data availability and its significance in influencing pipe failure behavior. Finding relevant covariates is an iterative process as characteristics considered relevant by the analyst or expert might not be adequately represented in the data. Different approaches have been used to determine covariates, covering non-parametric estimators (Carrión et al., 2010) and statistical tests (Fuchs-Hanusch et al., 2012; Kleiner and Rajani, 1999). The result of a statistical test is only valid in the context of the (failure) model the test was performed with. Cross-validation provides a more robust approach to assess the predictive uncertainty for a given set of covariates (e.g. Harrell, 2001, chapter 5).

The extension of a model with covariates is relatively straightforward. Covariates are usually taken into account by means of linear predictors $\boldsymbol{\beta}^T\mathbf{x}$, where $\boldsymbol{\beta}$ is a vector of additional parameters and $\mathbf{x}$ is the vector of covariates. Qualitative covariates (such as material) are considered with the help of dummy variables, see e.g. Montgomery et al. (2012).

The linear predictor can be incorporated into the model in different ways. One common way is the proportional hazards approach: the failure rate $\lambda_i(t)$ for pipe i with covariates $\mathbf{x}_i$ is

$$\lambda_0(t)\exp(\boldsymbol{\beta}^T\mathbf{x}_i)$$

where the base failure rate $\lambda_0(t)$ is defined by the model structure.[5] If covariates are included with the proportional hazards approach, it is theoretically

[5] The parameters $\boldsymbol{\beta}$ can be estimated without defining the base failure rate $\lambda_0(t)$ this is known as a Cox regression (e g Harrell 2001 chapter 19) and is an important reason for the popularity of proportional hazard models However to enable predictions of future failures the base failure rate must be specified

appealing to include the *logarithm* of the pipe length as covariate in $\mathbf{x}$. This results in the failure rate being directly scaled by the length l_i of a pipe segment, i.e. $\lambda_i(t) = \lambda_0(t)\, l_i^{\beta}$. If $\beta = 1$, then pipe segmentation has no influence; for example, the predicted number of failures is the same regardless of whether the predictions are made for a pipe with length l or split into two pipes with lengths l_1 and l_2 so that $l_1 + l_2 = l$.

3 2 Data requirements for calibration

All models have parameters that must be calibrated. In some cases, parameter values are available from other studies or experts are able to provide estimates. In most situations, however, it is preferable to calibrate the model with local failure data or data from national data bases (Grigg, 2009). The maximum likelihood (ML) method and Bayesian inference are two commonly applied approaches for data-based model calibration. The uncertainty of the estimated parameters can be quantified by both methods; Bayesian inference results in the complete distribution of the parameters, whereas confidence intervals can be approximated for ML estimates.

Both approaches require the formulation of the likelihood function: that is the conditional probability (density) of observing data $\mathbf{D}$ given a probabilistic model with parameters $\boldsymbol{\theta}$. It is often symbolized as $p(\mathbf{D}|\boldsymbol{\theta})$ or $\mathcal{L}(\boldsymbol{\theta})$. The formulation of the likelihood function is based on the model assumptions as well as on assumptions about the data quality and completeness. The latter point is particularly important, but sometimes overlooked. Pipe failure data typically show one or more of the following characteristics (Scheidegger et al., 2013):

- **Right censored observations:** As long as there are pipes in service, data is right censored. It corresponds to the time from the last failure or laying date to the time of observation.

- **Left truncation:** Left truncation occurs if a pipe was installed before failures were systematically recorded. As a consequence, it is not known how many failures occurred before the recording period and when (see Figure 3, ii and iv).

- **Absence of replaced pipe data:** Information about replaced pipes is often missing. Either because a pipe was replaced before data were recorded, or a replaced pipe was deleted from the database together with the corresponding pipe failure data because the database was established with the objective of reflecting the current state of the system.

This leads to "survival selection" (Renaud et al., 2011) due to the under-representation of pipes with poor failure histories in the data set which have already been removed (visualized in Figure 3, iii and iv).

These three characteristics result in the data situations depicted in Figure 3. If present in the data, any of these characteristics must be considered explicitly in the likelihood function. Ignoring this leads to systematically biased estimations that cannot be corrected by increasing the amount of data (e.g. Scheidegger et al., 2011). Thus, even under identical failure model assumptions, different likelihood functions must be derived depending on the data characteristics.

Theoretically, the likelihood function can be derived for any failure model and data situation. Practically, the mathematical expressions become elaborate and may lead to challenging numerical problems. For this reason Table 1 presents the data characteristics considered in the published likelihood functions of the models. Note that, given a likelihood function, both Bayesian inference and ML estimation are possible for any model. Bayesian inference and ML are alternative approaches for model calibration and not a property of the model.

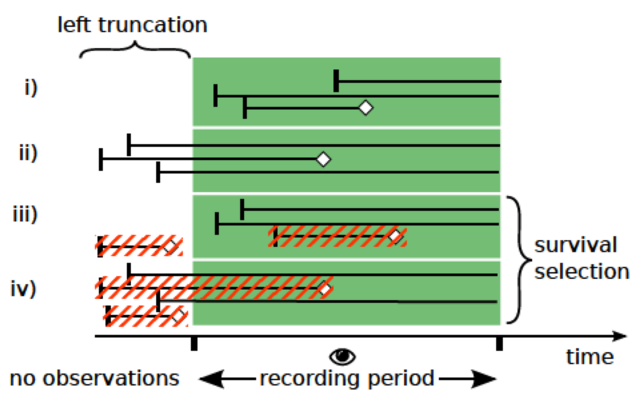


Figure 3 Common data situations for model calibration Records are only available for the duration of the recording period (green shaded) i) failures are recorded since the pipe laying date data of replaced pipes are available ii) only failures which occurred within the recording period are documented data of replaced pipes are available iii) failures are recorded since the laying date data of replaced pipes are unavailable (red hatched) iv) only failures within the recording period are documented data of replaced pipes are unavailable

3 3 Predictions

Prediction of future pipe failures, or more precisely the future failure probability, is often the main motivation for the use of pipe failure models. Statistical

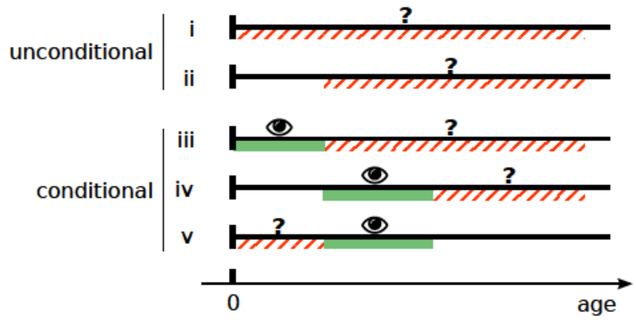


Figure 4 Different conditional and unconditional predictions For the hatched orange time periods failures should be predicted For the shaded green recording period failure records are available

pipe failure models treat failures as realizations of a random process. Therefore, only probabilistic statements about unobserved (future) failures can be made. Hence, every prediction must be expressed as probability distribution, hereafter referred to as *predictive distribution*. This distribution is usually summarized by statistics such as mean or standard deviation.

Different questions require different predictive distributions. Figure 4 illustrates five possible predictions of unobserved failures. If no information about the failure history of a pipe is available, only *unconditional* predictions are possible (Figure 4, i and ii). Unconditional predictions are useful to compare the failure behavior of different types of pipes, e.g. different materials and vintages. Mathematically, such a predictive probability distribution for the number of failures N in an interval I could be symbolized as

$$\text{Prob}(N_I).$$

It is common that some information of the pipe failure history is available. This should then be considered in the predictive distribution (Figure 4, iii–v). We denote this as *conditional* prediction. Conditional predictions are required to imulate how an existing pipe (or system) evolves over time. If H_J represents the known failure history of a pipe within the observation interval J, the predictive probability distribution of N_I is the conditional distribution

$$\text{Prob}(N_I \mid H_J) = \text{Prob}(N_I, H_J)/\text{Prob}(H_J).$$

If the predictions are formulated differently, for example as the distribution of the time to the next failure, the equations above must be rewritten correspondingly. However, the distinction between conditional and unconditional predictions still holds. Some models are based on the assumption that the probability for failures is independent of the failure history, i.e. $\text{Prob}(N_I \mid H_J) = \text{Prob}(N_I)$. Hence, in this situation conditional and unconditional predictions are identical.

Unconditional predictions are usually rather simple to formulate or to simulate with Monte Carlo procedures. The often more relevant conditional prediction can be cumbersome (although theoretically possible) to derive. Generally, the predictions become more complex the more the failure rate depends on the failure history of a pipe. For every model in this review, the kind of predictions described in the publications as analytic equations or Monte Carlo procedure are listed in Table 1.

4. Published models

The main properties of the reviewed statistical pipe failure models are summarized in Table 1, using the definitions presented in Section 3. For every model the failure rate λ is presented (sometimes slightly simplified to facilitate comparisons). The following notation is used throughout this section: pipe age is denoted by t, the pipe age at the ith failure by t_i with the convention $t_0 = 0$, the number of failures occurred before t by $n(t)$, and the kth model parameter as θ_k. A visualization of each failure rate can be found in Table 1.

Additional clarifications for each model are given when necessary, followed by a description of applications and software availability.

Eisenbeis (1994)

Model structure This model is based on the assumption of Weibull-distributed inter-arrival times between failures and can be expressed as failure rate

$$\lambda(t, t_n, n) = \begin{cases} \theta_1 \, (t)^{\theta_1 - 1}, & n(t) = 0 \\ \theta_2 \left(t - t_{n(t)}\right)^{\theta_2 - 1}, & 1 \leq n(t) < 4 \\ \theta_3 \left(t - t_{n(t)}\right)^{\theta_3 - 1}, & 4 \leq n(t) \end{cases}$$

Three different Weibull distributions are calibrated: for the time to the first failure, inter-arrival times up to the fourth failure, and the inter-arrival times of all following failures. Pipe properties and external conditions are considered by means of a proportional hazards approach. The effect of past failures is represented by an additional binary covariate stating whether any failure has occurred within 15 years since a previous failure.

The failure rate goes back to zero after each failure, hence implying that stress due to repairs is not modeled. Furthermore, continuous deterioration of the pipe cannot be considered.

Further developments The model is described in English in Eisenbeis et al. (1999), where the "number of previous failures" is incorporated as a covariate. Also, a detailed description of a Monte Carlo procedure for deriving the predictions by simulation is given. It takes explicitly into account that the number of previous failures varies in the course of the simulation. Another slight modification is described in Le Gat and Eisenbeis (2000). There, only a single Weibull distribution is estimated for all inter-arrival times. The problem of left-truncated data is discussed with the ad-hoc proposition of including a covariate that represents the age at the last known failure, if any, and otherwise the age at the beginning of observation period.

Model applications Of all models considered in this review, the model by Eisenbeis (1994) and its adaptations is perhaps the one with the largest number of reported applications (e.g. Alvisi and Franchini, 2010; Le Gat and Eisenbeis, 2000; Lei and Sægrov, 1998; Martins et al., 2013; Pelletier et al., 2003). The network sizes in the applications range from 155 to 1 243 km. All model applications use the pipe installation period or year as a model covariate.

Software The model by Eisenbeis et al. (1999) is implemented in CARE-W-PHM (Eisenbeis et al., 2002).

Gustafson and Clancy (1999)

Model structure This model is described as a Semi-Markov Process. Separate generalized Gamma distributions were calibrated to model the time between failures of different failure order. Then, statistical tests were performed to investigate whether the calibrated distributions are statistically different from exponential distributions. Although statistically significant differences were found, the authors have chosen to model all inter-arrival times after the first failures by means of exponential distributions:

$$\lambda(t, n) = \begin{cases} \theta_1 \theta_2 \left(\theta_2 t\right)^{\theta_1 - 1}, & n(t) = 0 \\ \theta_{n(t)+2}, & 1 \leq n(t) < 10 \\ \theta_{13}, & 11 \leq n(t) \end{cases}$$

The use of exponential distributions simplifies the implementation; however, no additional details are provided.

Exponential distributions result in a constant failure rate and hence cannot model deterioration over time. Still the model is quite flexible because separate rates up to the 11^{th} failure are estimated, which requires a large data set.

Calibration Gustafson and Clancy (1999) state that the predicted number of failures is unrealistically high with an "immediate and sustained increase of 50% [...] when compared to the previous five years." For comparison, a simpler model based only on exponential distributions was also calibrated, with a predicted increase of failures of only 20%. This difference is surprising, as the second model is a special case of the original model structure. The result may be explained by the reduction of the parameter uncertainty due to the smaller number of parameters of the simplified model.

Model applications This model was applied to at least two pipe systems in Canada; one of the systems is described in the paper that proposes the model, and the other is presented by Osman and Bainbridge (2011). In the first application pipe and failure data sets are not fully described but it is noteworthy that the first failure record dates back to 1958. In the second application the pipe and pipe failure data sets are also not described in detail, though the authors mention that the data used are left-truncated and right censored.

Software The distributions of the time between failures were estimated using the SAS `LIFEREG` procedure (SAS Institute, 1990).

Pelletier (2000)

Model structure In the paper proposing this model different combinations of Weibull-Exponential models were investigated. The most complex model assumes two independent Weibull distributions for the time to the first failures and inter-arrival time to the second failures, an exponential distribution for the inter-arrival time to the third failures, and another exponential distribution for all following inter-arrival times:

$$\lambda(t, t_n, n) = \begin{cases} \theta_1\theta_2\left(\theta_2 t\right)^{\theta_1 - 1}, & n(t) = 0 \\ \theta_3\theta_4\left(\theta_4(t - t_1)\right)^{\theta_3 - 1}, & n(t) = 1 \\ \theta_5, & n(t) = 2 \\ \theta_6, & n(t) \geq 3 \end{cases}$$

Maximum likelihood ratio tests were performed to compare the models. A shorter publication in English is available as well (Mailhot et al., 2000).

The model structure is somewhat similar to Gustafson and Clancy (1999) but simpler and hence with fewer parameters to estimate which may be advantageous for smaller data sets.

Calibration data This publication is probably the earliest in explicitly discussing the problem of left-truncated data in a rigorously mathematical way and in which a correct likelihood function for such data is derived.

Model applications Mailhot et al. (2000), Pelletier et al. (2003), Alvisi and Franchini (2010) and Toumbou et al. (2014) applied the model to five systems. The first two studies used the Chicoutimi (Canada) system, which has also been used in the application of the Eisenbeis (1994) model (e.g. Pelletier et al., 2003). The lengths of four of the five systems studied is shorter than 352 km. The study presented by Alvisi and Franchini (2010) applies the model to a large water distribution network with approximately 2 400 km of pipes.

Røstum (2000)

Model structure The model proposed by Røstum (2000) is based on the assumptions that the failure rate increases with the age of the pipe. The applied failure rate

$$\lambda(t) = \theta_1 \theta_2 t^{\theta_2 - 1}$$

corresponds to the Weibull distribution and covariates are included via a proportional hazards approach. This failure rate describes a non-homogeneous Poisson process (NHPP). An attractive feature of NHPP models is that the distribution of the number of failures in an interval can be expressed analytically (e.g. Cook and Lawless, 2010).

The main limitation of NHPP models is that while deterioration is represented, the influence of previous failures on the failure rate cannot be modeled.

Calibration A Poisson process is memoryless by definition, i.e. the failure rate is not influenced by the number of previous failures. This simplifies the calibration because left-truncated data do not require a different likelihood function.

Predictions Similarly, due to the memoryless property, all conditional predictions are mathematically identical to unconditional predictions.

Further developments Kleiner and Rajani (2010) and Economou et al. (2012) are modifications of the model of Røstum (2000). To discuss the implications of the modifications in detail, these models are separately presented in this review.

Model applications Røstum (2000) has applied this model to the 808 km long water distribution system of the city of Trondheim, Norway. The details of the water distribution system are well described in terms of earliest and latest installed pipe, total number of failures and earliest and latest recorded failure. The covariates used to demonstrate the applicability of the model were the logarithm of pipe length, the pipe diameter, time between construction, the start of failure recording and the logarithm of the number of previous failures.

Software The model by Røstum (2000) is implemented in CARE-W-NHPP ("Winroc", Eisenbeis et al., 2002; Sægrov et al., 2003) and PARMS (Burn et al., 2003).

Watson et al (2004)

Model structure The model is formulated as a homogeneous Poisson process, hence the failure rate

$$\lambda = \theta_1$$

is a constant. This is the simplest possible model structure which is likely to be inadequate for most data sets.

Calibration Bayesian inference to calibrate the rate parameters is proposed. The posterior distribution is derived analytically for a gamma distributed prior for the failure rate. An extension for a hierarchical Bayesian model is discussed but no details are provided.

Predictions As a homogeneous Poisson process is memoryless, the conditional predictions are identical to the unconditional ones, hence the model can not make use of information about the failure history of a pipe.

Model applications No application with real data is known. Watson et al. (2004) demonstrated the model with artificially generated data.

Economou et al (2008)

Model structure This model can be seen as an extension of Røstum (2000). While the used failure rate

$$\lambda(t) = \theta_1 t^{\theta_1 - 1}$$

is similar, zero-inflation is modeled additionally and for every pipe, three individual parameters are introduced (referred to as random effects). As a result, in total $3n + 3$ parameters have to be estimated, with n being the number of pipes in the calibration data. These parameters cannot be identified from the data alone; Bayesian inference with an informative prior distribution is required. Given the very large number of parameters, it can be expected that the influence of the prior distribution is considerable. Unfortunately, the selection of the prior distribution is not discussed and the chosen values appear somewhat arbitrary. Moreover, the authors do not elaborate much on why three random effects per pipe are required and no tests are presented to establish whether these are beneficial for the prediction. The zero inflation models a system in which some pipes fail according to a Poisson process while other pipes do not fail at all. As a result, the distribution of the number of failures of a pipe in a given interval is a mixture of a Poisson distribution and a probability mass for zero failures.

The failure rate of the Poisson process reflects the deterioration of pipes. However, it is very difficult to reason about how the zero-inflation effects the failure rate, especially as the probability that a pipe can fail depends on the age and other covariates. Furthermore, the presented equation how the zero-inflation is combined with the failure model is erroneous or at least unclear: the expressed likelihood function seems to contain random variables (not the probability density of random variables).

Predictions Similar to the model of Røstum (2000), the conditional predictions reduce to unconditional predictions. However, a slight complication is that the probability that a pipe can fail is implemented as a function of covariates including the "age at the end of the observation period", assuming that older pipes are more likely to fail. It is not mentioned in the description of the Monte Carlo procedure used for predictions whether this age covariate is changed within the prediction horizon (as in future the pipe will be older).

Model applications Application of this model has been reported in at least two publications (Economou et al., 2012; Kleiner and Rajani, 2012).

The systems are relatively small, i.e. less than 1 500 pipes (approximately 150 km pipe length for the case study presented in Kleiner and Rajani (2012)), located in the USA, New Zealand and Canada. The pipe and pipe failure data sets are reasonably described. Contrary to the most common covariate used in the other models applications, installation year or period have not been used in this model applications. Other model covariates such as pipe length, diameter, maximum absolute pressure and pressure change were included.

Software The model was implemented in WinBUGS (Lunn et al., 2000) a software to perform Bayesian inference in an automated way.

Le Gat (2009)

Model structure The model is based on an extension of the Yule process (Athreya and Ney, 2004) which enables the likelihood functions and predictive distributions to be expressed analytically as negative binomial distributions.

The model structure can simultaneously capture the deterioration of pipes representing an increase of the failure rate after each failure:

$$\lambda(t, n) = (1 + \theta_1 n(t))\ \theta_2 t^{\theta_2 - 1}$$

Le Gat (2009) provides a complete description of the model which can be seen as the gold standard for model documentation. Recently, a shorter description of the model has been published in English in Le Gat (2014).

Calibration data Likelihood functions for different data characteristics are derived. The problem of survival selection is discussed and a model extension to correct for it proposed. It is based on the assumption that the probability of a pipe not to be replaced after a failure is a function of pipe age.

Model applications This model has been applied to several real cases (e.g. Claudio et al., 2014; Le Gat, 2014; Martins et al., 2013). The length of the water distribution systems used to demonstrate the applicability of the models range from 367 to 3 081 km. Similar to other applications, the water distribution systems are not fully described in terms of earliest installed pipe, latest installed pipe, total number of failures and earliest failure record. The covariates used in all three applications of the model are the logarithm of pipe length and the pipe diameter. Except in Martins et al. (2013), additional pipe characteristics such as pipe location, bedding conditions, connection type, and climate data were used.

Software This model has been implemented in at least two software packages: Casses (Renaud et al., 2012)[6] and AWARE-P (Cardoso et al., 2012)[7].

Kleiner and Rajani (2010)

Model structure The model is very similar to the one of Røstum (2000). It is presented as a non-homogeneous Poisson process (NHPP), but differs slightly in the assumption that the failure rate is piece-wise constant with a step size of one year:

$$\lambda(t) = \theta_1 \lfloor t \rfloor^{\theta_2}$$

The main innovation is that Kleiner and Rajani (2010) introduced "time-dependent covariates" to account for seasonal influences on pipe breaks.

If the model is interpreted as NHPP, only deterioration is considered. This is not the case if the number of previous failures enters the model as covariate. However, as discussed in the next paragraph that is a violation of the NHPP assumptions.

Calibration The problem of left truncation is recognized but mathematically not treated in a formal way. Instead, a covariate representing the "number of *known* previous failures" is introduced. This variable is classified as "pipe and time-dependent", which suggests that the failure rate is regarded by the authors as a function of age and the number of previous failures, $\lambda(t, n)$. In this case, however, in contrast to the introduction of Kleiner and Rajani (2010) the model is not memoryless anymore and hence does not describe a Poisson process (compare e.g. Cook and Lawless, 2010, chapter 2.2).

Predictions As a result of including the "number of *known* previous failures" the standard equations for NHPP models must not be applied to compute predictions. This complicates the computation of predictive distributions so that, for example, a Monte Carlo procedure similar to the one proposed by Eisenbeis et al. (1999) would be necessary. However, a discussion of this problem is missing in Kleiner and Rajani (2010).

[6] http://casses.irstea.fr/en/ archived http://www.webcitation.org/6UB9CZ6bw
[7] http://www.aware-p.org/ archived http://www.webcitation.org/6UB8xroUh

Note, that the other time-dependent covariates, such as a freezing index, do not contradict the Poisson process assumptions. They are included to describe the climate influences. While "climate-related covariates can be used to train the model on observed historical breaks but not to forecast" (Kleiner et al., 2010), they may prevent unusual climate events from influencing the failure rate estimation and make the parameter estimates more comparable between different systems.

Further developments Kleiner and Rajani (2010) also present an extension of the model to account for zero-inflation. It is similar to the approach proposed by Economou et al. (2008), with the difference that no random effects are estimated for individual pipes.

Model applications Three water distribution systems were used to test this model. One of the systems was presented in the original paper (i.e. Kleiner and Rajani, 2010), and two other applications were presented by Kleiner et al. (2010) and (Osman and Bainbridge, 2011). The first two model applications were performed in a small system, 1 091 pipes representing 146.6 km of network and 490 pipes representing 66 km of pipe network. The third application was conducted using a system with more than 30 320 pipes. The covariates used were pipe material, pipe length, diameter, installation data and climate data.

Software I-WARP (Kleiner and Rajani, 2010). An earlier, multi-variate exponential pipe failure time model by Kleiner and Rajani (2004) is the basis for I-WARP s predecessor D-WARP.

Scheidegger et al (2013)

Model structure This is a model based on the assumptions that the time to the first failure can be modeled as Weibull distribution and all the subsequent times between failures with one exponential distribution. The corresponding failure rate is

$$\lambda(t,n) = \begin{cases} \theta_1\theta_2\left(\theta_2 t\right)^{\theta_1 - 1}, & n(t) = 0 \\ \theta_3, & n(t) \geq 1 \end{cases}$$

These are strong assumptions resulting in an inflexible failure rate that cannot represent deterioration over time and is only partly influenced by the previous failures. However, these simplifications reduce the number of parameters and hence facilitate the elicitation of prior distributions which may be important for small data sets.

Calibration The model was presented to exemplify a general blueprint on how the likelihood function of any failure model can be adapted to data lacking the information on replaced pipes so as to avoid survival selection bias (see Section 3.2). This is achieved by combining the failure model with a replacement model that encapsulates the assumptions of the past replacement activities. This approach is similar to the one by Le Gat (2014), but the probability of a pipe not being replaced depends on the number of previous failures and not directly on the pipe age.

Predictions Scheidegger et al. (2013) presents the equations for predictions of type i–vi (illustrated in Figure 4). The conditional predictions of type v are described in the appendix of Scholten et al. (2014).

Model applications This model was demonstrated in Scheidegger et al. (2013) using data of a part of the water distribution system of Lausanne, Switzerland, consisting of about 3 000 pipe segments. Scholten et al. (2014) applied the model to four mid-size to small Swiss water utilities (60–715 km) with data affected by left truncation and selective survival to support long-term asset management and also discussed practical limitations.

Software The model is implemented in R (R Development Core Team, 2012) and freely available on request.

5. Discussion

This paper presents a review of state-of-the-art statistical pipe failure models as defined in Section 2. Focusing on this class of deterioration models enables a more detailed comparison based on three intuitive but often neglected points: i) clarification of model assumptions; ii) description of the assumed data characteristics; and iii) presentation of the formulated probabilistic predictions. The main contribution of this review is that the different models are presented in a comparable way. This involves reformulating all models as failure rates and making implicit assumptions about data structures and predictions transparent.

This systematic approach provides guidance in selecting a suitable model for a given data situation and desired type of prediction. For example, assume a water utility with data characterized by a short observation window and no records of replaced pipes (i.e. left-censoring and selective survival, see Figure 3-iv). The aim is to predict the future number of failures of the pipes. The failure histories for these pipes is partly known and should be considered (prediction type iv in Fig. 4). In this case, Table 1 shows that all required equations or Monte Carlo procedures are only available for two models: Le Gat (2009) and Scheidegger et al. (2013). If instead, the data situation was described by case ii in Figure 4 (left truncation, no selective survival), the models of Pelletier (2000), Watson et al. (2004), and Economou et al. (2008) could also be correctly calibrated. The respective functions for a type iv-prediction would still need to be derived in the case of the Pelletier (2000) model.

To further discriminate between these models, the description of the failure rate can be considered as it quickly reveals the fundamental model assumptions, such as pipe deterioration or the influence of the number of failures. We derived the failure rate for every model in order to enable a comparison among the differently mathematically formulated models. For example, the failure rate of the model by Gustafson and Clancy (1999) offers higher flexibility regarding the representation of subsequent failures than the models of Scheidegger et al. (2013) or Pelletier (2000). But the later models require fewer parameters to be estimated and could be the better choice in situations of small networks with only a few pipe records (Scholten et al., 2014).

The uncertainty of the prediction is influenced by various factors such as the model assumptions, the amount of calibration data, the qualitative characteristics of the data, the fraction of pipes with observed failures, and how much relevant information is contained in the covariates. Due to this complexity, simple statements on "how much data" a certain model requires are not valid. Model selection is about finding a compromise between model flexibility and the number of parameters that is adequate for a given data set.

The comparison of the models presented in this review was not a straightforward task. This was partly because some of the important references are not available in English, and partly because all model assumptions are not always mathematically clearly defined, which makes a full understanding and replication of the models difficult. Furthermore, the model predictions are frequently neglected or poorly documented in the available publications. This is surprising as predictions are the main motivation to apply a pipe failure model.

For all models we briefly described their applications as published in scientific journals. The data of the pipe systems studied are not always completely characterized; for example, it is often unclear if the data are left truncated or if data of replaced pipes are included or not. It is therefore not possible to corroborate what kind of data most of studied networks have and whether the available models are adequate.

The review also identified potential for model improvements. New research could be conducted in two main directions: i) Making existing models more widely applicable by deriving additional likelihood functions and predictive distributions. Table 1 shows many opportunities in this direction; especially extensions to avoid survival selection bias do not yet exist for most of the reviewed models and conditional predictions for more complex models are often missing. ii) Developing new models (or extending existing ones) to mimic failure behavior more realistically. The conceptual failure rate presented in Figure 2 gives some guidance. Furthermore, one aspect that is not yet addressed by any model is the spatial dependency of failures, i.e. the increased chance of failures of pipes located in the vicinity of failed pipes.

A major hindrance for practical application is a lack of easy-to-use software tools for model calibration and failure prediction. The information regarding the capabilities of the different software is patchy and the model structure, data assumptions, and possible predictions are often not clearly stated. However, even the most user-friendly software requires an at least intuitive understanding of the underlying model to judge how trustworthy the obtained results are.

Developing good strategies for water distribution and other urban water infrastructures is crucial to ensure long-term service reliability and good quality. Pipe failure models alone cannot provide the solution but are definitely a key part of it.

6. Conclusions

- Pipe failure models have a significant advantage when compared to pipe life span models because the failure behavior is modeled independent of management decisions.
- Left-truncation and survival selection are common characteristics of the data available in urban water utilities. However, these characteristics are not taken into account by all models. Different characteristics of the calibration data require differently formulated likelihood functions, i.e. models shall be selected or developed based on the given data characteristics.
- For many failure models presented in this review the equations or methods to compute pipe failure predictions are sparsely documented; this is especially true for conditional predictions.

- Most of the available pipe failure models consider either an increase of the failure rate with pipe age or the failure rate is influenced by the previous pipe failures. From the models herein reviewed, there is not a single model structure that appears superior.

- The selection of the most adequate pipe failure model needs to be guided taking into account the following points: (i) the appropriate likelihood function should be selected based on the data characteristics; (ii) the model predictions should match the initial questions, and (iii) the assumption of the failure rate should be in agreement with the experience of the operator.

- Although there is a relatively large number of pipe failure models available there is still room for further development. One of the features the reviewed models still do not account for is the potential spatial dependency of failures—a problem frequently reported by urban water distribution network managers.

7. Acknowledgments

We thank Max Maurer an the two anonymous reviewers for their constructive inputs, which greatly improved this manuscript.

8. References

References

Alegre, H., Baptista, J., Cabrera, E., Cubillo, F., Duarte, W., Hirner, W., Merkel, W., Parena, R., 2006. Performance Indicators for Water Supply Services, 2nd Edition. IWA International Water Association.

Alvisi, S., Franchini, M., 2010. Comparative analysis of two probabilistic pipe breakage models applied to a real water distribution system. Civil Engineering and Environmental Systems 27 (1), 1–22.

Athreya, K. B., Ney, P. E., 2004. Branching Processes. Dover Publications, Mineola, New York.

Burn, S., Tucker, S., Rahilly, M., Davis, P., Jarrett, R., Po, M., 2003. Asset planning for water reticulation systems – the PARMS model. Water Science and Technology: Water Supply 3 (1-2), 55–62.

Cardoso, M. A., Santos Silva, M., Coelho, S. T., Almeida, M. C., Covas, D. I. C., 2012. Urban water infrastructure asset management – a structured approach in four water utilities. Water Science and Technology 66 (12), 2702–2711.

Carrión, A., Solano, H., Gamiz, M. L., Debón, A., 2010. Evaluation of the reliability of a water supply network from right-censored and left-truncated break data. Water resources management 24 (12), 2917–2935.

Claudio, K., Couallier, V., Le Gat, Y., 2014. Integration of time-dependent covariates in recurrent events modelling: application to failures on drinking water networks. Journal de la Socit Franaise de Statistique 155 (3), 62–77.

Cook, R. J., Lawless, J. F., Apr. 2010. The Statistical Analysis of Recurrent Events, 1st Edition. Springer, New York.

Economou, T., Kapelan, K., Bailey, T. C., 2012. On the prediction of underground water pipe failures: zero inflation and pipe-specific effects. Journal of Hydroinformatics 14 (4), 872–883.

Economou, T., Kapelan, Z., Bailey, T., 2008. A zero-inflated Bayesian model for the prediction of water pipe bursts. In: Proceedings of Water Distribution Systems Analysis 2008.

Eisenbeis, P., 1994. Modélisation statistique de la prévision des défaillances des conduites d eau potable. Ph.D. thesis, Université Luise Pastor.

Eisenbeis, P., Le Gat, Y., Laffréchine, K., Le Gauffre, P., König, A., Rostum, J., Tuhovcak, L., Valkovic, P., 2002. CARE-W WP2 description and validation of technical tools: D3 – models description. Tech. rep., European Commission Fifth Framework Programme Technical Report No. EVK1-CT-2000-00053.

Eisenbeis, P., Røstum, J., Le Gat, Y., 1999. Statistical models for assessing the technical state of water networks – Some European experiences. In: Proceedings of the AWWA Annual Conference.

Fuchs-Hanusch, D., Kornberger, B., Friedl, F., Scheucher, R., 2012. Whole of life cost calculations for water supply pipes. Water Asset Management International 8, 19–24.

Grigg, N. S., 2009. National mains failure database project. In: Proceedings of ASCE-EWRI World Environmental and Water Resources Congress, Kansas City, MO.

Gustafson, J. M., Clancy, D. V., 1999. Modeling the occurrence of breaks in cast iron water mains using methods of survival analysis. In: Proceedings of the Annual Conference of the AWWA, Chicago.

Harrell, F. E., 2001. Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer, New York.

Kleiner, Y., Nafi, A., Rajani, B., Dec. 2010. Planning renewal of water mains while considering deterioration, economies of scale and adjacent infrastructure. Water Science & Technology: Water Supply 10 (6), 897–906.

Kleiner, Y., Rajani, B., 1999. Using limited data to assess future needs. American Water Works Association 91 (7), 47–61.

Kleiner, Y., Rajani, B., 2001. Comprehensive review of structural deterioration of water mains: statistical models. Urban Water 3 (3), 131–150.

Kleiner, Y., Rajani, B., 2004. Quantifying effectiveness of cathodic protection in water mains: theory. Journal of Infrastructure Systems 10 (2), 43–51.

Kleiner, Y., Rajani, B., 2012. Comparison of four models to rank failure likelihood of individual pipes. Journal of Hydroinformatics 14 (3), 659–681.

Kleiner, Y., Rajani, B. B., 2010. I-WARP: Individual water main renewal planner. Drinking Water Engineering and Science 3, 71–77.

Le Gat, Y., 2009. Une extension du processus de Yule pour la modélisation stochastique des événements récurrents. Application aux défaillances de canalisations d eau sous pression. Ph.D. thesis, Cemagref Bordeaux, Paristech.

Le Gat, Y., 2014. Extending the Yule process to model recurrent pipe failures in water supply networks. Urban Water Journal 11 (8), 617–630.

Le Gat, Y., Eisenbeis, P., 2000. Using maintenance records to forecast failures in water networks. Urban Water 2 (3), 173–181.

Le Gat, Y., Kropp, I., Poulton, M., 2013. Is the service life of water distribution pipelines linked to their failure rate? Water Science and Technology: Water Supply 13 (2), 386–393.

Lei, J., Sægrov, S., 1998. Statistical approach for describing failures and lifetimes of water mains. Water Science and Technology 38 (6), 209–217.

Liu, Z., Kleiner, Y., Rajani, B., Wang, L., Condit, W., 2012. Condition assessment technologies for water transmission and distribution systems. Tech. rep., U.S. Environmental Protection Agency.

Lunn, D. J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and computing 10 (4), 325–337.

Mailhot, A., Pelletier, G., Noël, J., Villeneuve, J., 2000. Modeling the evolution of the structural state of water pipe networks with brief recorded pipe break histories: Methodology and application. Water Resources Research 36 (10), 3053–3062.

Martins, A., Leito, J., Amado, C., 2013. Comparative study of three stochastic models for prediction of pipe failures in water supply systems. Journal of Infrastructure Systems 19 (4), 442–450.

Montgomery, D. C., Peck, E. A., Vining, G. G., 2012. Introduction to Linear Regression Analysis, 5th Edition. Wiley, Hoboken, New Jersy.

Nishiyama, M., Filion, Y., 2013. Review of statistical water main break prediction models. Canadian Journal of Civil Engineering (40), 972–979.

OECD, 2006. Infrastructure to 2030: Telecom, Land transport, Water and Electricity. OECD Publishing.

Osman, H., Bainbridge, K., 2011. Comparison of statistical deterioration models for water distribution networks. Journal of Performance of Constructed Facilities 25, 259–266.

Pelletier, G., 2000. Impact du remplacement des conduites d aqueduc sur le nombre annuel de bris. Ph.D. thesis, Université du Québec, Institut national de la recherche scientifique.

Pelletier, G., Mailhot, A., Villeneuve, J.-P., 2003. Modeling water pipe breaks – three case studies. Journal of Water Resources Planning and Management 129, 115–123.

R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

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

Renaud, E., Le Gat, Y., Poulton, M., 2011. Using a break prediction model for drinking water networks asset management: From research to practice. In: Leading Edge conference on Strategic Asset Management (LESAM). Mühlheim an der Ruhr, Germany.

Renaud, E., Le Gat, Y., Poulton, M., 2012. Using a break prediction model for drinking water networks asset management: From research to practice. Water Science and Technology: Water Supply 12 (5), 674–682.

Røstum, J., 2000. Statistical modelling of pipe failures in water networks. Ph.D. thesis, Norwegian University of Science and Technology NTNU.

Saegrov, S., 2005. CARE-W: computer aided rehabilitation of water networks. IWA Publishing.

Sægrov, S., Schilling, W., Røstum, J., Tuhovcak, L., Eisenbeis, P., Herz, R., Le Gauffre, P., Melo Baptista, J., Conroy, P., de Federico, V., Mazzacane, S., Schiatti, M., 2003. Computer-aided rehabilitation of water networks (CARE-W). Water Science and Technology: Water Supply 3 (1), 19–27.

SAS Institute, 1990. SAS/STAT user s guide. Version 6. SAS Institute Cary, NC.

Scheidegger, A., Hug, T., Rieckermann, J., Maurer, M., 2011. Network condition simulator for benchmarking sewer deterioration models. Water Research 45 (16), 4983–4994.

Scheidegger, A., Scholten, L., Maurer, M., Reichert, P., 2013. Extension of pipe failure models to consider the absence of data from replaced pipes. Water Research 47 (11), 3696–3705.

Scholten, L., Scheidegger, A., Reichert, P., Mauer, M., 2013. Combining expert knowledge and local data for improved service life modeling of water supply networks. Environmental Modelling & Software 42, 1–16.

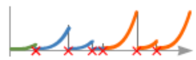
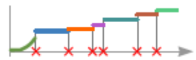
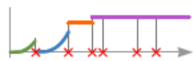
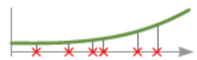
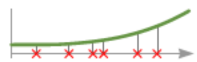
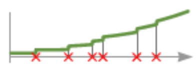
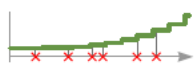
Scholten, L., Scheidegger, A., Reichert, P., Mauer, M., Lienert, J., 2014. Strategic rehabilitation planning of piped water networks using multi-criteria decision analysis. Water research 49, 124–143.

St. Clair, A. M., Sinha, S., Jan. 2012. State-of-the-technology review on water pipe condition, deterioration and failure rate prediction models! Urban Water Journal 9 (2), 85–112.

Toumbou, B., Villeneuve, J., Beardsell, G., Duchesne, S., 2014. General model for water-distribution pipe breaks: Development, methodology, and application to a small city in Quebec, Canada. Journal of Pipeline Systems Engineering and Practice 5 (1).

Watson, T. G., Christian, C. D., Mason, A. J., Smith, M. H., Meyer, R., 2004. Bayesian-based pipe failure model. Journal of Hydroinformatics 6 (4), 259–264.

Table 1: Important properties of statistical pipe failure models. The mathematical dependencies of the failure rate λ are shown with t denoting the pipe age, t_i the age at the ith failure, and n the number of previous failures. Different colors in the graphical representation of the failure rate separate parts with independent parameterization.

Publication	Math formulation	Failure rate		spatial dependencies	Covariates	Calibration (cf. Fig. 3)				Predictions (cf. Fig. 4)				
						i)	ii)	iii)	iv)	i)	ii)	iii)	iv)	v)
Eisenbeis (1994)	time	$\lambda(t, t_n, n)$ [†]		–	✓	✓	–	–	–	✓[†]	✓[†]	–	–	–
Gustafson and Clancy (1999)	time	$\lambda(t, n)$ [‡]		–	✓	✓	–	–	–	✓	✓	✓	–	–
Pelletier (2000)	time	$\lambda(t, t_n, n)$		–	–	✓	✓	–	–	✓	✓	–	–	–
Røstum (2000)	count	$\lambda(t)$		–	✓	✓	✓[*]	–	–	✓	✓	✓[*]	✓[*]	✓[*]
Watson et al. (2004)	count	$\lambda(\cdot)$		–	–	✓	✓[*]	–	–	✓	✓[*]	✓[*]	✓[*]	✓[*]
Economou et al. (2008)	count	$\lambda(t)$ [§]		–	✓	✓	✓[*]	–	–	✓	✓	✓[*]	✓[*]	✓[*]
Le Gat (2009)	count	$\lambda(t, n)$		–	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Kleiner and Rajani (2010)	count	$\lambda(t)$ [§‖]		–	✓	✓[*‖]	–	–	–	✓	✓[*‖]	✓[*‖]	✓[*‖]	✓[*‖]
Scheidegger et al. (2013)	time	$\lambda(t, n)$		–	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

[*] Not mentioned in the publication. However, the conditional predictions reduce mathematically to the unconditional predictions. Similarly, left-truncation of the data has no influence on the calibration.
[†] Based on the modification according to Eisenbeis et al. (1999).
[‡] Referring to the more complex model in the first part of the publication.
[§] Failure rate without influence of the zero-inflation.
[‖] Only if based on the interpretation as Poisson model.