Water Resour Manage (2011) 25:2509–2524
DOI 10.1007/s11269-011-9823-0

# Estimating the Timing of the Economical Replacement of Water Mains Based on the Predicted Pipe Break Times Using the Proportional Hazards Models

**Suwan Park**

Received: 3 August 2010 / Accepted: 4 April 2011 /
Published online: 19 April 2011


**Abstract** Deteriorated water mains fail frequently causing service disruption and other inconvenience to the customers. Therefore, the utilities must conduct repair, rehabilitation and/or replacement in a timely manner to satisfy the needs of the customers. To succeed in this process the utilities must also consider the economics of the water main maintenance. The proposed methodology presents not only a method for the optimal maintenance but also a practical way of conducting it by providing the economical time period of maintenance. A method is also presented for analyzing the accuracy of proportional hazards models (PHMs) in forecasting break times and estimating the timing for economical replacement of water mains. A survival probability criterion for the forecasting of the pipe breaks was determined in order to minimize the prediction errors of the PHMs. Subsequently, the criterion was used to estimate the upper and lower bounds of future break times of a water main using the survival functions derived from the PHMs. Two General Pipe Break Prediction Models (GPBMs) for a pipe were estimated for each of the two series of the recorded and predicted, upper and lower bound break times. The threshold break rate (TBR) was coupled with the two GPBMs for each pipe and solved for time to give the upper and lower bounds of the economical replacement time period.

**Keywords** Break prediction model · Economical replacement · Pipe break · Prediction error · Proportional hazards model · Survival function · Water mains

S. Park (✉)
Department of Civil and Environmental Engineering, Pusan National University,
Busan 609-735, South Korea
e-mail: swanpark@pusan.ac.kr

## 1 Introduction

Water mains represent one of the most fundamentally important infrastructure assets. Therefore, it is crucial for a water authority or utility to maintain the water mains in their most desirable state by conducting timely maintenance. Determining the times for timely maintenance requires an understanding of the break characteristics of the water mains and the uncertainty that inherently exists in the assessment of the characteristics. This uncertainty in our understanding of the pipe break mechanism necessitates the development of more practical guidance for the efficient maintenance of water mains.

The previous methodologies to find the optimal time of pipe replacement were focused on the rather precise timing of replacement, e.g., elapsed years or months since installation. However, the models and/or the parameters that were estimated to model pipe break patterns during the processes of the methodologies inescapably contain uncertainties. Lack of attention to these uncertainties led to estimating the time of replacement which also bears uncertainty. Moreover, the utilities may not find the exact time of replacement very useful because of other circumstances such as road pavement that they may have to consider in establishing the pipe replacement schedule.

This paper presents a methodology to assess an optimal time period for the replacement of water mains considering the economical aspects of pipe maintenance and the predictive power of the Proportional Hazards Model (PHM) (Cox 1972) used for the pipe break pattern modeling. The methodology was built upon the survival function form of the PHMs that were developed by Park et al. (2011) in which seven PHMs had been developed for the consecutive break times of the case study water pipes.

In this paper, a method to analyze the precision of the PHMs in predicting the pipe break times was also developed. Using the case study water pipes in Park et al. (2011) the survival probability that minimizes the prediction errors of the recorded break times was determined, and used in predicting the subsequent future break times of the pipes. A procedure for estimating the economical replacement time period of a water main was subsequently illustrated based on the analysis results of the PHM accuracy.

## 2 Previous Studies on Efficient Water Main Maintenance

Studies on the development of a methodology to determine the economical timing of water main maintenance originated from the work by Shamir and Howard (1979) who used an economic decision model in conjunction with an exponential model to predict the number of breaks for a given time. A similar method was applied by Clark et al. (1982) to estimate the optimal replacement time of a pipe. Walski and Pelliccia (1982) used the concept of critical break rate to identify "bad" pipes that need to be replaced based on the comparison between the current break rate and a critical break rate. Kleiner and Rajani (2001) provided a comprehensive literature review on the break pattern models along with the methodologies related to optimal maintenance of water mains.

Loganathan et al. (2002) also used the concept of critical break rate, called the threshold break rate (TBR), which they developed by applying a discrete cost discounting scheme. However, Walski and Pelliccia (1982) used a continuous discounting scheme to develop a critical break rate requiring a specified planning period and an exponential break prediction model.

Loganathan et al. (2002) improved on the critical break rate of Walski and Pelliccia (1982) by obtaining the TBR without using a planning period or a break prediction model. In addition, they proved that the TBR and the break rate model are equivalent by definition, which supported a methodology to estimate the economically optimal timing of water main replacement.

Park and Loganathan (2002) developed a pipe break prediction model that can model any break pattern ranging from a linear to an exponential trend by introducing a weighting factor in the model. This new model has been termed the General Pipe Break Prediction Model (GPBM). Coupled with the TBR of Loganathan et al. (2002), the GPBM is subsequently used in estimating the economical replacement time of water mains (Park and Loganathan 2002).

Dandy and Engelhardt (2001) suggested the use of genetic algorithms to provide a rehabilitation strategy for a deteriorating water distribution system and also reviewed other categories of economic pipe maintenance methodology such as the application of optimization techniques (de Schaetzen et al. 1998; Halhal et al. 1997; Kim and Mays 1994; Lansey et al. 1992), some variants of the economic models (Male et al. 1990; Walski and Pelliccia 1982), and system-wide decision models (Deb et al. 1998; Kleiner et al. 1998).

Park et al. (2007) used the PHMs constructed for case study water mains to establish the time intervals of optimal water main replacement. Since the methodology of estimating the optimal replacement time interval was based on the predicted break times using the constructed PHMs, it is crucial to analyze the prediction accuracy of the constructed PHMs. However, in their method, Park et al. (2007) did not take into account the accuracy of the constructed PHMs in predicting the pipe break times.

The upper and lower limits of the predicted break times were simply assumed in Park et al. (2007) using two arbitrary survival probabilities while the two limits were estimated in this paper based on the confidence intervals (CIs) of the estimated survival functions at the best survival probability obtained through the analyses on the predictive power of the PHMs. In addition, the PHMs used in Park et al. (2007) did not include the time-dependent covariates while the PHMs utilized in this paper are the models developed in Park et al. (2011) in which the time-dependent covariates were modeled.

Some other works related to efficient management of water pipe networks are as follows. Pinto et al. (2010) proposed a theory of vulnerability of water pipe networks to identify the most vulnerable parts of a water pipe network based on a clustering process for pipelines by establishing hierarchical description of the network. Carrión et al. (2010) used non- and semi-parametric methods with a modified Nelson–Aalen estimator to analyze the pipe failure probabilities and characteristics, such as influencing factors for pipe failure, of case study pipes of which left-truncation and right-censoring failure data were present. Christodoulou and Deligianni (2010) developed a neuro-fuzzy decision-support system that can perform multi-factored risk-of-failure analysis and aid pipe asset management of a water distribution network. Christodoulou et al. (2010) presented a framework that utilized a combination of artificial neural network and both parametric and nonparametric survival analysis techniques to investigate the identified risk factors of pipe failure and estimate the forecasted time to failure metric.

## 3 Overview of the Proposed Methodology

The water main break database used for this study is from a mid-western state of the U.S.A. The database contained information on the pipe material type, joint type, installation time, break time, diameter and length. Additionally, the database contained information on the locations where pipe material/joint type change. The pipes were also managed using the utility's proprietary grid management system that stored information regarding internal pipe pressure type, degree of land development, and number of customers for each grid.

Park et al. (2011) provides more detailed descriptions on the dataset used for the construction of the PHMs that were utilized for the analyses in this study. The individual pipes that were defined according to the intrinsic characteristics of the pipes and the information used for the grid system in Park et al. (2011) were also used for the analyses in this paper. Park et al. (2011) also extensively discussed about the procedures of constructing the PHMs and the corresponding survival functions with detailed explanations regarding the explanatory variable (factor) selection processes.

As illustrated in Fig. 1, the methodology for estimating the economically optimal replacement time period of a water main begins with constructing the PHMs for consecutive pipe breaks. The PHMs utilized for this purpose were the ones constructed by Park et al. (2011). Equations 1, 2, 3, 4, 5, 6 and 7 represent the survival functions that were derived from the PHMs in Park et al. (2011) using the analytical relationship between a hazard function and survival function.

$$\hat{S}(t) = \exp\left\{-\exp\left(\frac{28.867}{1+e^{-2.613\ln t+14.87}} - 23.172\right)\right\}^{\exp\left(\begin{array}{l}1.3424\,TYPE+0.5447\,DL+0.0005\,L+0.0197\,C+0.000259\,DL\cdot L\\-0.0128\,TYPE\cdot C-0.0230\,TYPE\cdot time-0.00003\,C\cdot time\end{array}\right)} \quad (1)$$

$$\hat{S}(t) = \exp\left\{-\exp(-7.187)\cdot t^{1.272}\right\}^{\exp\left(\begin{array}{l}0.1785\,TYPE+0.3444\,DL+0.1704\,L+1.2570\,C\\-0.5310\,DL\cdot C-1.6262\,C\cdot time\end{array}\right)} \quad (2)$$

$$\hat{S}(t) = \exp\left\{-\exp(-3.361)\cdot t^{0.600}\right\}^{\exp(0.2565\,SR-0.0600\,SF+0.8549\,L-0.1700\,C)} \quad (3)$$

$$\hat{S}(t) = \exp\left\{-\exp(-2.927)\cdot t^{0.609}\right\}^{\exp(0.7291\,L)} \quad (4)$$

$$\hat{S}(t) = \exp\left\{-\exp(-2.417)\cdot t^{0.548}\right\}^{\exp(-0.8784\,SR-1.0185\,SF+0.6473\,DL-0.3247\,L)} \quad (5)$$

$$\hat{S}(t) = \exp\left\{-\exp(-2.642)\cdot t^{0.627}\right\}^{\exp(0.4415\,L)} \quad (6)$$

$$\hat{S}(t) = \exp\left\{-\exp(-2.072)\cdot t^{0.551}\right\}^{\exp(0.6686\,L)} \quad (7)$$

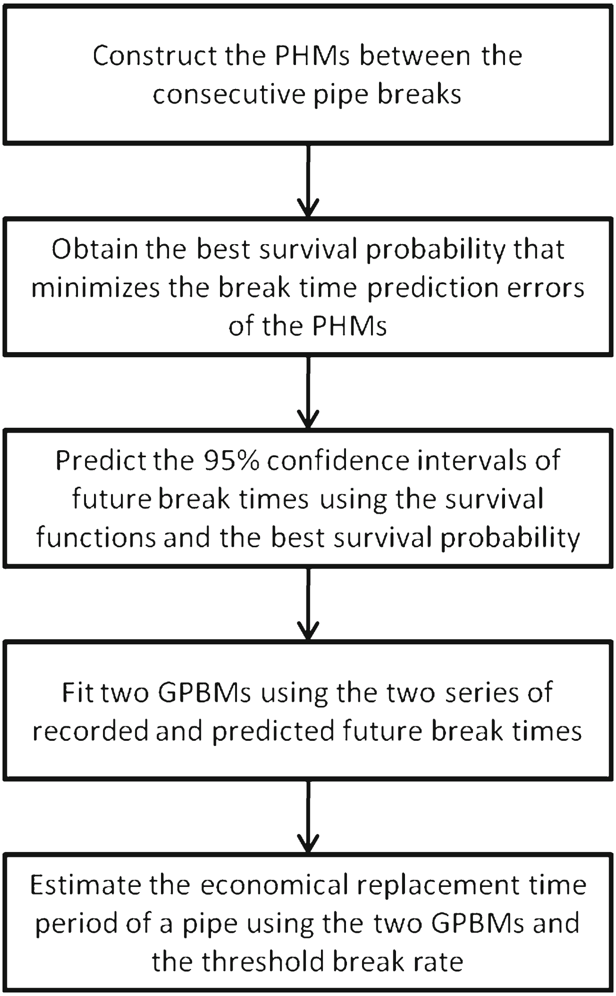



**Fig. 1** Procedures of estimating the economical replacement time period of a pipe

In Eqs. 1–7, *SR*, *SF*, *TYPE*, *DL*, *L*, and *C* represent spun cast iron pipe with a rigid joint, spun cast iron pipe with a flexible joint, pipe material type, degree of land development, length, and number of customers, respectively. These variables take on either binary or continuous values depending on the corresponding characteristics of a pipe. Detailed descriptions on the values of these explanatory variables and the implications of them in the hazard analysis of pipes were explained in detail in Park et al. (2011).

Each of Eqs. 1–7 represents the survival function for the Survival Time Group (STG) I $\sim$ VII in Park et al. (2011), respectively. These STGs were used in Park et al. (2011) to group the case study pipes according to the number of recorded number of breaks. For example, STG II is a group of the pipes that have at least 1 recorded break. The survival function for STG II is modeled based on the elapsed time between the 1st recorded break time and the 2nd recorded break time or the time when the observation ended.

Using a survival function one can obtain the time that corresponds to a specific survival probability. Therefore, in this paper, the break times of the pipes in the break data used in Park et al. (2011) were predicted using the survival functions with a given survival probability, and compared with the recorded break times of the pipes in Park et al. (2011). Since the precision of the prediction will depend on the survival probability that is used to obtain the predicted break times, many values of the survival probability were utilized in comparing the predicted and the recorded break times. The prediction precision was analyzed for each set of recorded and predicted break times for each STG and the survival probability used. Then, the survival probability that produced least errors for the most of the STGs in predicting the actual break times was selected as the best survival probability.

Since even the predicted future break times using this best survival probability will bear uncertainty, the extent of this uncertainty was estimated based on the 95% CIs of the estimated survival functions to provide the upper and lower bounds of the predicted break times.

Therefore, two series' of break times were predicted for a pipe and, then, each series of the predicted break times was fitted with the GPBM by Park and Loganathan (2002). Coupled with the threshold break rate by Loganathan et al. (2002) the two limits of the economically optimal replacement time were estimated using each of the predicted break time series. The time interval between the two estimated economical replacement times was determined as the economical replacement time period of a pipe of interest. The following sections provide detailed illustrations of the methodology and the analysis results.

## 4 Assessment of the Predictive Power of the PHMs

### 4.1 Successive and Cumulative Break Prediction Accuracy of the PHMs

Two analysis methods were used. The first assesses the accuracy of individually predicted break times using the PHMs in Park et al. (2011) for each order of break occurrence. This accuracy was analyzed using the Individual Break Prediction Error (*IBPE*) which was estimated as:

$$IBPE_n = R_n - P_n \tag{8}$$

where $n$ represents the order of breaks of a pipe which ranges from 1 to 7, $R_n$ the elapsed time between the recorded $(n-1)$th and $n$th break times, and $P_n$ the elapsed time between the recorded $(n-1)$th break time and the predicted $n$th break time that is estimated using the survival function of STG $n$. That is, the survival functions estimated as from Eqs. 1–7 were used to predict a consecutive break time, which is $P_n$, of a pipe.

For example, $IBPE_3$ represents the error of the predicted 3rd break time elapsed from the recorded 2nd break time using the survival function for STG III. That is, it is the difference between the time from the recorded 2nd break to the recorded 3rd break and the time from the recorded 2nd and the predicted 3rd break. $IBPE_3$ is conceptually depicted in Fig. 2, in which the difference between the times $R_3$ and $P_3$ is $IBPE_3$. In Fig. 2, $O_n$ is the recorded $n$th break time and $X_n$ is the predicted $n$th break time.

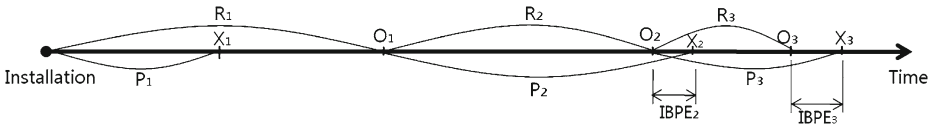



**Fig. 2** Calculation concept of $IBPE_3$

The second method of analyzing the predictive power of the PHMs assesses the accuracy of the predicted *n*th break time which is continuously predicted from the 1st break time up to the *n*th break time. This accuracy was analyzed using the Consecutive (or Cumulative) Break Prediction Error (*CBPE*) which was estimated as:

$$CBPE_n = \sum_{i=1}^{n} R_i - \sum_{i=1}^{n} P_i \tag{9}$$

where $n$ represents the order of breaks of a pipe which ranges from 1 to 7, $R_i$ the elapsed time between the recorded $(i-1)$th and $i$th break times, and $P_i$ the elapsed time between the recorded $(i-1)$th break time and the predicted $i$th break time that is estimated using the survival function of STG $i$.

For example, $CBPE_3$ is the difference between the recorded 3rd break time and the cumulatively predicted 3rd break time which is estimated by accumulating the predicted 1st, 2nd, and 3rd break times using the survival functions of STG I, II, and III. Therefore, $CBPE_3$ represents the difference between the time from installation to the recorded 3rd break and summation of the time from installation to the predicted 1st break time estimated using the survival function of STG I, the time from the predicted 1st break to the predicted 2nd break estimated using the survival function of STG II, and the time from the predicted 2nd break to the predicted 3rd break estimated using the survival function of STG III.

Figure 3 depicts the calculation concept of $CBPE_3$. Therefore, $CBPE_3$ is calculated as the difference between $(R_1 + R_2 + R_3)$ and $(P_1 + P_2 + P_3)$ in Fig. 3. $P_2$ in Fig. 2 and $P'_2$ in Fig. 3, are the same in size. Similarly, $P_3$ in Fig. 2 and $P'_3$ in Fig. 3 are the same in size.

## 4.2 Analysis of the IBPE and CBPE

The survival function provides the time that corresponds to a specified survival probability. Therefore, a survival probability may be selected and the corresponding time may be used as the predicted break time. The usual practice in the reliability engineering and medical statistics field is to select a survival probability of '0.5', called

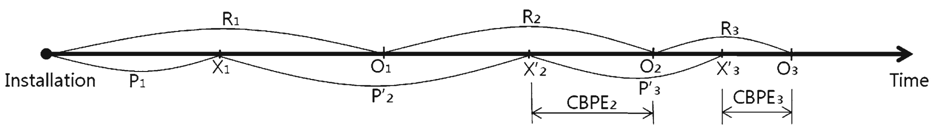



**Fig. 3** Calculation concept of $CBPE_3$

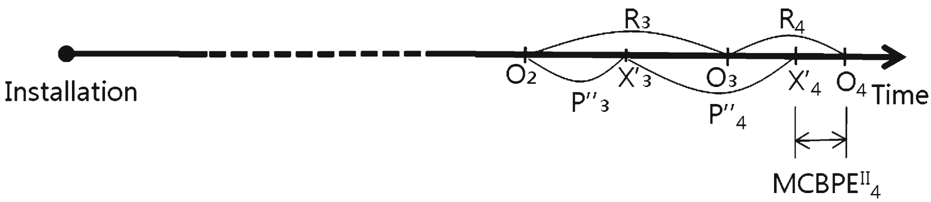



**Fig. 4** Calculation concept of $MCBPE_4^{II}$

the median survival time, to estimate the representative time of a lifetime of an item or a person (Fig. 4).

However, since the purpose of the present analysis is to find the best survival probability that minimizes the overall break time prediction errors, it is inappropriate to use only the median survival times to predict pipe break times. The best survival probability may be obtained by analyzing the prediction errors for various values of the survival probability. Therefore, a reasonable range of survival probabilities was used to calculate the *IBPE*s and *CBPE*s.

With a given survival probability and the corresponding covariate values of the pipes the break times of the pipes in Park et al. (2011) were predicted using the survival functions of the STGs. Tables 1 and 2 present the absolute maximum values of the *IBPE*s and *CBPE*s, respectively, that determine 95% of the area of the relative frequency histograms of the *IBPE*s and *CBPE*s with respect to the value of '0' for the survival probabilities of 0.60, 0.65, 0.70, and 0.75, respectively. The zero value represents the case of zero error in the predictions. The survival probabilities other than 0.60, 0.65, 0.70, and 0.75 resulted in larger *IBPE*s and *CBPE*s than the ones shown in Tables 1 and 2.

Table 1 presents the prediction errors for successive break times that are the times of the next order of break after recorded break times. For example, for the case of STG V, the maximum error of predicting the fifth break time was estimated as 3.97 years when the survival probability of 0.70 for the survival function of STG V is used to predict the fifth break times.

Based on Table 1, it was conjectured that the successive break time of a pipe from the 3rd break time and beyond could be predicted using the constructed PHMs with a reasonable error, given that predicting the successive break times of a water main with great precision is extremely difficult. On the other hand, large *IBPE* values have been computed for STG I and STG II. Therefore, using Models I and II in predicting the first and second break times of a pipe was expected to result in large errors. Moreover, the large errors that occurred in predicting the first and second break

**Table 1** Absolute maximum of the *IBPE*s (years)

| Survival probability | STG I | STG II | STG III | STG IV | STG V | STG VI | STG VII |
|---|---|---|---|---|---|---|---|
| 0.60 | 35.7 | 42.3 | 8.76 | 3.05 | 4.27 | 2.89 | 2.42 |
| 0.65 | 35.6 | 38.5 | 6.78 | 3.53 | 3.99 | 2.74 | 2.51 |
| 0.70 | 35.6 | 34.6 | 6.13 | 4.19 | 3.97 | 3.02 | 2.55 |
| 0.75 | 35.6 | 30.7 | 6.52 | 4.06 | 4.28 | 3.32 | 2.73 |

**Table 2** Absolute maximum of the *CBPE*s (years)

| Survival probability | STG I | STG II | STG III | STG IV | STG V | STG VI | STG VII |
|---|---|---|---|---|---|---|---|
| 0.60 | 35.7 | 38.9 | 37.9 | 38.7 | 41.0 | 36.9 | 40.3 |
| 0.65 | 35.7 | 36.8 | 35.3 | 36.0 | 36.7 | 33.9 | 36.0 |
| 0.70 | 35.6 | 35.7 | 33.0 | 32.3 | 31.4 | 32.3 | 34.3 |
| 0.75 | 35.6 | 34.3 | 30.6 | 31.4 | 29.6 | 32.2 | 35.9 |

times accumulated in the process of continuously predicting the 3rd break time and beyond, thereby increasing the *CBPE*s, as shown in Table 2.

Due to the large errors in predicting the 1st and 2nd break and the effects of the errors on the successive predictions, it was conjectured that the continuously predicted 3rd or more break times for the pipes with '0' or '1' recorded breaks will be unreliable.

## 4.3 Modified CBPE

Since there existed large errors in cumulatively predicting the break times of the pipes starting from the 1st break, the prediction errors for the 3rd break time and beyond based on the recorded second break time were estimated by $MCBPE_n^{II}$(Modified *CBPE* based on the 2nd break time) which is expressed as:

$$MCBPE_n^{II} = \sum_{i=3}^{n} R_i - \sum_{i=3}^{n} P_i \tag{10}$$

where $n$ stands for the order of the predicted break time and ranges from 3 to 7.

For example, $MCBPE_4^{II}$, which represents the cumulative error of predicting the 4th break time starting from the recorded second break time, was calculated as the difference between $(R_3 + R_4)$ and $(P_3'' + P_4'')$ as shown in Fig 3.

Table 3 presents the absolute maximum values of the 95% of the $MCBPE_n^{II}$ histogram for STG III and beyond for each of the selected survival probabilities.

A comparison of Tables 2 and 3 revealed that the prediction errors for the case of predicting the 3rd break and beyond starting from the recorded 2nd break time were much less than the case of predicting the successive break times starting from the 1st break time. Therefore, it was concluded that the survival functions of STG III through VII were appropriate in continuously predicting the consecutive break times of the water mains in the case study area, provided that at least two breaks have occurred on a pipe.

In other words, the continuously predicted break times of the pipes with '0' or '1' recorded breaks were not considered to be reliable due to the cumulative effects of

**Table 3** Absolute maximum values of the $MCBPE_n^{II}$ (years)

| Survival probability | STG III | STG IV | STG V | STG VI | STG VII |
|---|---|---|---|---|---|
| 0.60 | 8.75 | 10.8 | 11.5 | 12.9 | 13.7 |
| 0.65 | 6.7 | 7.7 | 8.5 | 8.8 | 9.4 |
| 0.70 | 6.1 | 6.8 | 8.3 | 8.1 | 10.4 |
| 0.75 | 6.5 | 8.6 | 10.7 | 11.0 | 13.0 |

the large prediction errors for the predicted 1st and 2nd break times. In the mean time, the continuously predicted 3rd to 7th break times for the pipes that had at least two breaks were considered to be reliable because the large errors in predicting the 1st and 2nd break times did not apply in the first place.

Furthermore, as can be observed from the results in Table 3, except for the 7th break predicting the break times using the survival probability of 0.70 consistently produced smaller errors than the other survival probabilities used. Therefore, the survival probability of 0.70 was selected to predict the break times of the water mains under study.

## 5 Estimating the Timing for Economical Replacement of Water Mains

### 5.1 Procedures to Estimate the Economical Replacement Timing

The methodology to estimate the economical replacement timing of an individual pipe developed in this paper utilizes the 95% confidence interval (CI) of the consecutively predicted break times using the survival functions of the PHMs of the case study pipes in Park et al. (2011) and the GPBM developed by Park and Loganathan (2002).

The "baseline" statement in the SAS system provides estimates of the upper and lower bounds of the baseline survival probabilities calculated for each of the recorded break times of the STGs. An appropriate function was fitted for each series of the upper and lower bounds of the baseline survival probabilities to obtain the upper and lower bound baseline survival functions of a pipe.

Then, the exponential covariate term in the PHM estimated for an STG was multiplied to the fitted baseline survival functions to obtain the upper and lower bound survival functions of a pipe. The upper and lower bounds of the predicted break times were estimated as the corresponding times of the fitted upper and lower bound survival functions for the survival probability of 0.7.

The upper and lower bounds of the 95% CI of the consecutively predicted break times were used as additional break time data in the GPBM, which is expressed as:

$$N_c(t) = (1 - WF) \cdot (Bl + Al \cdot t) + WF \cdot \exp(Ae \cdot t + Be) \tag{11}$$

where $N_c(t)$ is the cumulative number of breaks at time $t$, $WF$ the weighting factor, and $Ae$, $Be$, $Al$, and $Bl$ the curve fitting coefficients.

The GPBM is capable of fitting any pipe break trends, irrespective of whether they are exponential, linear, or in between the two, using the weighting factor, $WF$. The regression coefficients and the $WF$ of the GPBM were estimated using the least squares method and the algorithm developed by Park and Loganathan (2002). Since a pipe had two series' of break times, two GPBMs were obtained for each pipe. That is, one of the GPBMs utilized the recorded break times and the lower bounds of the 95% CI of the predicted break times, and the other one utilized the recorded break times and the upper bounds of the 95% CI of the predicted break times.

The economical replacement timing of a pipe was obtained by considering the upper and lower bounds of the optimal replacement times estimated using the two GPBMs constructed for a pipe. The upper and lower bounds of the optimal replacement time period were estimated based on the equivalence relationship between the TBR (Loganathan et al. 2002) and the GPBM, which is expressed as:

$$\frac{dN_c(t)}{dt} = \frac{\ln\{(1+R)/(1+i)\}}{\ln(1+C/F)} \tag{12}$$

where $t$ is the break time in elapsed years since installation, $R$ the interest rate(1/yr), $i$ the inflation rate(1/yr), $C$ the repair cost, and $F$ the replacement cost. The left and right hand sides of Eq. 12 represent the time derivative of the GPBM and the TBR, respectively.

The TBR represents the break rate of a pipe for which the present worth of the total costs of repair and replacement is at a minimum. The upper and lower bounds of the economical replacement time period of an individual pipe as elapsed years from installation were obtained by solving Eq. 12 with respect to time using the two GPBMs. The equation to obtain the lower or upper bound of the economical replacement time period is expressed as:

$$t^* = \frac{\ln\left(\dfrac{TBR + WF\cdot Al - Al}{Ae\cdot WF}\right) - BE}{Ae} \tag{13}$$

## 5.2 An Example of Estimating the Economical Replacement Timing

An example to illustrate the methodology of estimating the economical replacement time period of a water main is provided below using the pipe break records and the covariate values shown in Table 4. The replacement cost per unit length of this 812 m-long pipe is US \$300, while the repair cost per break is US \$3,120. The annual interest and inflation rates used for the calculation of the TBR are 4.5% and 1%, respectively.

*SR*, *SF*, *L*, and *C* in Table 4 represent spun cast iron rigid joint pipe, spun cast iron flexible joint pipe, pipe length, and customers in a grid, respectively. Since this pipe had only two recorded break times, subsequent future break times could be predicted using the developed methodology. For example, the upper and lower bounds of the predicted 3rd break time were estimated using the 95% CI of the survival functions of probabilities estimated at each recorded break time of the pipes in STG III.

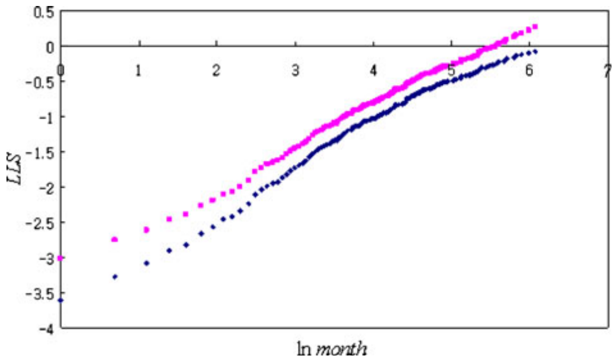
Figure 5 shows the plot of the upper and lower bounds of the 95% CI of the *LLS* (log-log-transformed values of survival probabilities or ln(-ln $S_0(t)$)) with respect to the log-transformed values of the recorded break times (ln $t$) for STG III.

The upper and lower bound baseline survival functions were obtained by fitting a linear function for each series of the *LLS* estimates shown in Fig. 5 and transforming

**Table 4** Specifications of the example pipe

| *SR* | *SF* | *DL* | *L* | *C* | Installation time | Recorded break times (years) |
|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 1 | 1 | Jan. 1953 | 1996.0, 1997.0 |

**Fig. 5** Upper and lower bounds of the 95% CI of the baseline survival probabilities for STG III




them into the survival functions. Then, the upper and lower bound survival functions of STG III were obtained as:

$$S_U = \exp\left\{-e^{-3.1828}\, t^{0.5851}\right\}^{\exp(0.2565\, SR - 0.0600\, SF + 0.8549\, L - 0.1700\, C)} \tag{14}$$

$$S_L = \exp\left\{-e^{-3.5818}\, t^{0.6205}\right\}^{\exp(0.2565\, SR - 0.0600\, SF + 0.8549\, L - 0.1700\, C)} \tag{15}$$

where $S_U$(t) and $S_L$(t) represent the upper and lower bound survival functions for STG III, respectively.

The upper and lower bounds of the predicted 3rd break time of a pipe in STG III were estimated by setting Eqs. 14 and 15 to 0.7 and solving them for time. Therefore, by using the covariate values in Table 4, the upper and lower bounds of the predicted 3rd break time as elapsed time since the recorded 2nd recorded break time were computed as:

$$\begin{aligned} t_{3U} &= \left(-\frac{\ln 0.70 \cdot e^{3.1828}}{e^{0.2565\times 0.3698 - 0.06\times(-0.3359) + 0.8549\times 0.2280 - 0.1699\times 0.5562}}\right)^{1/0.5851} \\ &= 27.4\ \text{months from the 2nd break} = \text{year } 1999.3 \end{aligned} \tag{16}$$

$$\begin{aligned} t_{3L} &= \left(-\frac{\ln 0.70 \cdot e^{3.5818}}{e^{0.2565\times 0.3698 - 0.06\times(-0.3359) + 0.8549\times 0.2280 - 0.1700\times 0.5562}}\right)^{1/0.5851} \\ &= 43.1\ \text{months from the 2nd break} = \text{year } 2000.6 \end{aligned} \tag{17}$$

where $t_{3U}$ and $t_{3L}$ represent the upper and lower bounds of the predicted 3rd break time, respectively.

Using the same procedure, the 95% CI values of the predicted break times of the example pipe up to the 7th break were obtained, as shown in Table 5.

**Table 5** The 95% CI values of the predicted break times

| Bound | 3rd break | 4th break | 5th break | 6th break | 7th break |
|---|---|---|---|---|---|
| Upper bound (year) | 1999.3 | 2000.3 | 2000.9 | 2001.6 | 2001.9 |
| Lower bound (year) | 2000.6 | 2002.1 | 2003.1 | 2004.4 | 2005.0 |

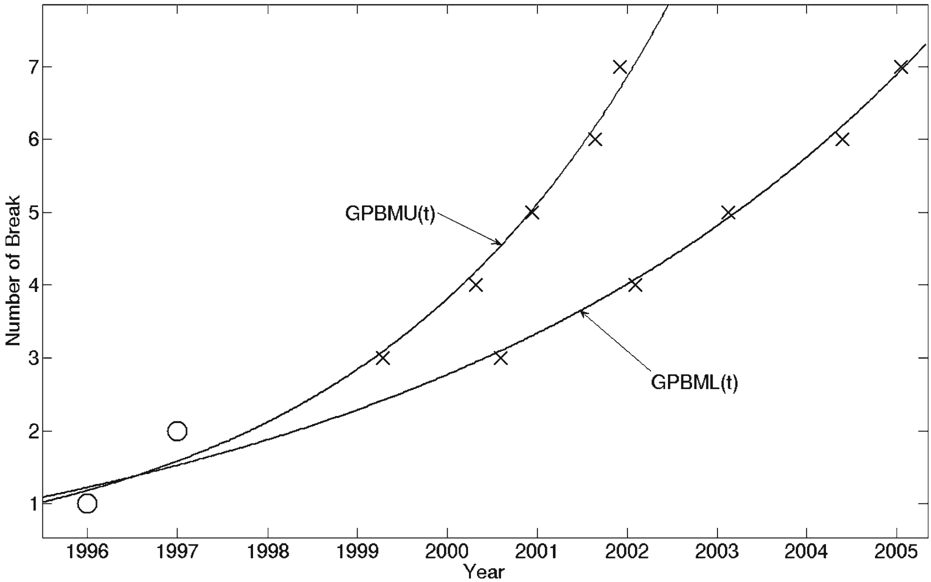



**Fig. 6** Graphs of the GPBMU and GPBML for the example pipe

The GPBMs utilizing the recorded break times and the upper and lower bounds of the predicted break times of the example pipe in Table 5 were obtained based on Park and Loganathan (2002) as:

$$\mathrm{GPBMU}\,(t) = \exp\,(0.2933 \cdot t - 585.2) \tag{18}$$

$$\mathrm{GPBML}\,(t) = 0.12 \cdot (0.5964 \cdot t - 1189.4) + 0.88 \cdot \exp\,(0.1898 \cdot t - 378.5) \tag{19}$$

where GPBMU and GPBML are the GPBMs utilizing the upper and lower bounds of the predicted break times, respectively.

Figure 6 represents Eqs. 18 and 19 graphically, in which 'o' and 'x' represent the recorded and predicted break times, respectively.

The economical replacement time period of the example pipe was obtained as [2003.0, 2009.2] by substituting the parameters in Eqs. 18 and 19 into Eq. 13.

This method of estimating the economical replacement timing of a pipe was applied to the case study water mains. Tables 6 and 7 summarize the statistics of the economical replacement time periods and the exact times for economical replacement elapsed from the installation times, respectively. The exact time point for economical replacement was obtained by predicting the consecutive break times using the survival probability of 0.7 for each survival function of the STGs and using

**Table 6** Summary statistics of the economical replacement time periods (years)

| Category | Recorded number of breaks | | | | |
|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 |
| Mean | 7.3 | 4.5 | 1.9 | 1.3 | 0.4 |
| Median | 5.4 | 2.8 | 1.3 | 0.9 | 0.2 |
| Standard deviation | 4.4 | 4.5 | 1.6 | 1.0 | 0.4 |

**Table 7** Summary statistics of the exact times for economical replacement (years since installation)

| Category | Recorded number of breaks | | | | |
|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 |
| Mean | 58.7 | 65.6 | 58.2 | 61.3 | 62.02 |
| Median | 55.2 | 59.3 | 54.8 | 58.2 | 60.46 |
| Standard deviation | 19.9 | 28.2 | 22.3 | 19.0 | 11.27 |

them in conjunction with the GPBM. Therefore, in this case, the 95% CI of the predicted break times was not considered.

Table 6 reveals that the economical replacement time periods decrease with increasing recorded number of breaks. This phenomenon is considered to be the result of the fact that the gaps between the upper and lower bounds of the predicted break times tend to decrease as there is less number of breaks to be predicted.

## 6 Summary and Conclusions

This paper has presented a methodology for analyzing the accuracy of the PHMs in predicting the future break times of water mains. In addition, the analysis results were used to estimate the time period for economical replacement of water mains. The survival functions of the PHMs constructed for a water main break database were used to predict the break times. The survival probability that resulted in the least prediction errors for the most of the STGs was determined as 0.7, which used as the criterion to estimate the future break time of the pipes.

To estimate the economical time period for replacement, two GPBMs for a pipe were estimated for the two series of recorded and predicted break times. The predicted break times used to create the two series of break times were obtained by calculating the times for the upper and lower bound survival functions of the successive break orders that correspond to the survival probability of 0.7.

The upper and lower bound survival functions of the pipes were obtained by estimating the upper and lower bounds of the 95% CI of the baseline survival probabilities of the STGs and multiplying them with the exponential covariate function in the estimated PHMs. The TBR was coupled with the two GPBMs for each pipe and solved for time to give the upper and lower bounds of the economical replacement time period for a pipe.

The economical replacement time period statistics for the case study pipes showed that the time period decreased with increasing number of recorded breaks, while the statistics for the exact economical replacement times did not reveal any apparent pattern with regard to the recorded number of breaks.

Deteriorated water mains physically and functionally fail frequently causing service disruption and other inconvenience to the customers such as traffic congestion and loss of revenue. Moreover, it is generally perceived that deteriorated water mains are the main cause of the degradation of customer confidence on the water supply systems. Therefore, the utilities must conduct repairs, rehabilitation and/or replacement in a timely manner to satisfy the needs of the customers.

To succeed in this process the utilities must also consider the economics of the water main maintenance. The proposed methodology presented not only a method for the economical maintenance but also a practical way of conducting it by providing the economical period of maintenance. The methodology took into account the uncertainty in predicting the pipe break times using the survival functions derived from the PHMs and suggested an analysis procedure to minimize the prediction errors.

Furthermore, it utilized this uncertainty to provide a reasonable time period for economical pipe replacement, thereby providing a margin of safety for the economical maintenance of water mains. The utilities could benefit from the proposed methodology in terms of having leeway in the implementation of the maintenance schedule with confidence on the economic optimality of the maintenance.

The methodology is based on the assumptions that a considerably large number of breaks have been recorded for a water distribution system, that the break data have been maintained for a relatively long period of time, in the order of several decades, and that the PHMs for the recorded consecutive break times can be constructed. Once the data are collected, organized and used to construct the PHMs, the managers of water distribution systems will benefit from this methodology by gaining an increased level of confidence in planning the long-term maintenance and in allocating the funds for the economical replacement of the pipes in their systems.

**Acknowledgements** This work was supported by the Financial Supporting Project of Long-term Overseas Dispatch of PNU's Tenure-track Faculty, 2010.

## References

Carrión A, Solano H, Gamiz ML, Debón A (2010) Evaluation of the reliability of a water supply network from right-censored and left-truncated break data. Water Resour Manage 24:2917–2935

Christodoulou S, Deligianni A (2010) A neurofuzzy decision framework for the management of water distribution networks. Water Resour Manage 24:139–156

Christodoulou S, Agathokleous A, Charalambous B, Adamou A (2010) Proactive risk-based integrity assessment of water distribution networks. Water Resour Manage 24:3715–3730

Clark RM, Stafford CL, Goodrich JA (1982) Water distribution systems: a spatial and cost evaluation. J Water Resour Plan Manage ASCE 108(3):243–256

Cox DR (1972) Regression models and life tables. J R Stat Soc 34(B):187–220

Dandy GC, Engelhardt M (2001) Optimal scheduling of water pipe replacement using genetic algorithms. J Water Resour Plan Manage ASCE 127(4):214–223

de Schaetzen W, Randall-Smith MJ, Savic D, Walters GA (1998) A Genetic algorithm approach for rehabilitation in water supply systems. In: Proc. Int. Conf. on Rehabilitation Technol. for Water Industry, pp 1–11

Deb AK, Hasit YJ, Grablutz FM, Herz RK (1998) Quantifying future rehabilitation and replacement needs of water mains. AWWA Research Foundation

Halhal D, Walters GA, Ouzar D, Savic DA (1997) Water network rehabilitation with a structured messy genetic algorithm. J Water Resour Plan Manage ASCE 123(3):137–146

Kim JH, Mays LW (1994) Optimal rehabilitation model for water-distribution systems. J Water Resour Plan Manage ASCE 120(5):674–692

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

Kleiner Y, Adams BJ, Rogers JS (1998) Selection and scheduling of rehabilitation alternatives for water distribution systems. Water Resour Res 34(8):2053–2061

Lansey KE, Basnet C, Mays LW, Woodburn J (1992) Optimal maintenance scheduling for water distribution systems. Civ Eng Syst, E & FN Spon, Lond 9:211–256

Loganathan GV, Park S, Sherali HD (2002) A threshold break rate for pipeline replacement in water distribution systems. J Water Resour Plan Manage ASCE 128(4):271–279

Male JW, Walski TM, Slutsky AH (1990) Analyzing water main replacement policies. J Water Resour Plan Manage ASCE 116(3):362–374

Park S, Loganathan GV (2002) Optimal pipe replacement analysis with a new pipe break prediction model. J Korean Soc Water Wastewater, KSWW 16(6):710–716
Park S, Kim JW, Agbenowosi N, Jun H (2007) A methodology to estimate economically optimal replacement time interval of water distribution pipes. Water Sci Technol Water Supply 7(5):149–155
Park S, Jun H, Agbenowosi N, Kim BJ, Lim K (2011) The proportional hazards modeling of water main failure data incorporating the time-dependent effects of covariates. Water Resour Manage 25(1):1–9
Pinto J, Varum H, Bentes I (2010) A theory of vulnerability of water pipe network (TVWPN). Water Resour Manage 24:4237–4254
SAS Institute Inc. (2009). http://www.sas.com/index.html
Shamir U, Howard CDD (1979) An analytic approach to scheduling pipe replacement. J Am Water Works Assn 71(5):248–258
Walski TM, Pelliccia A (1982) Economic analysis of water main breaks. J Am Water Works Assn 74(3):140–147