NRC Publications Archive Archives des publications du CNRC

Using limited data to assess future needs

Kleiner, Y.; Rajani, B. B.

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. /
La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version
acceptée du manuscrit ou la version de l'éditeur.

Publisher's version / Version de l'éditeur:

Journal (American Water Works Association), 91, July 7, pp. 47-61, 1999-07-01

NRC Publications Record / Notice d'Archives des publications de CNRC:

<https://nrc-publications.canada.ca/eng/view/object/?id=66879e90-d62b-4b6a-bdc1-09018a51e6b1>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=66879e90-d62b-4b6a-bdc1-09018a51e6bb>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

National Research
Council Canada

Conseil national
de recherches Canada

NRC - CNRC

Using limited data to assess future needs

Kleiner, Y. ; Rajani, B. B.

NRCC-42499

Reprinted from / Réimprimé par
Journal (American Water Works Association), v. 91, no. 7, July 1999, pp. 47-62, by
permission American Water Works Association.

For additional information relevant to the drinking water industry, visit
<http://www.awwa.org>

www.nrc.ca/irc/ircpubs

Using limited data to assess future needs

A five-step methodology was developed to help utilities with sketchy or scarce data on water main break history to budget for future pipe replacement.

**Yehuda Kleiner
and Balvant Rajani**

Repair and replacement of aging water mains impose major expenditures on municipalities and regional governments. The need for these expenditures is increasing as old pipes continue to deteriorate. Studies have been performed

Most major urban utilities are aware of the need to keep records of distribution system water mains. These records—which typically include data about water main break history, pipe material, pipe vintage, and soil type—are eventually used to allocate budgets to replace, rehabilitate, or repair aging distribution systems. Many utilities have incomplete or scarce water main data, which may make financial planning difficult. Thus, a methodology that facilitates the use of scarce, available data to assess future needs is warranted. Presented is a case study that followed five essential steps: (1) bundle data on water main breaks into homogeneous groups to predict future breaks for water utility regions for which data were insufficient, (2) establish breakage rate patterns for these groups to project future breakage rates, (3) use projected breakage rates to determine the economic life of water mains, (4) examine probabilistic scenarios of water main life, and (5) determine the investments required to replace water mains.

on the subject of water main failure. Whereas some studies investigated the physical mechanisms of pipe breakage,¹⁻⁶ others attempted to identify statistical relationships between water main break rates and influential factors such as age of pipes, soil corrosivity, pipe diameter, prior pipe breaks, temperature, operating pressure, and external loads.⁷⁻¹⁸ A short description of some of the

**For executive summary,
see page 171.**

Standard Deviations for Probability Distributions of Pipe-replacement Timing

Uncertainty in pipe-replacement timing should reflect three factors

The uncertainty in the probability distribution of pipe-replacement timing should reflect (1) the uncertainty associated with the parameter derivation from the 24 years of data from region Z, (2) the uncertainty inherent in using these parameters to predict breakage rates of another (presumably similar) population, and (3) the variability of individual segments within a group.

Thus, the standard deviations SD_i for the Gumbel and Weibull probability distributions of pipe-replacement timing in group i were heuristically calculated as follows:

$$SD_i = S_i \times L_i \times D_i \times E_i \quad (1S)$$

in which

$$S_i = \sqrt{\frac{\sum_{t=1}^{24} \left[\frac{N(t)_i^u - L_i^u \times N(t_0)_i \times e^{\alpha_i(t-t_0)}}{24-2} \right]^2}{24-2}} \quad (2S)$$

$$L_i = \sqrt{\frac{L_i^u}{L_i^r}} \quad (3S)$$

$$D_i = 1 + \frac{\left| \sum_{t=21}^{24} \frac{N(t)_i^u}{L_i^u} - \sum_{t=21}^{24} \frac{N(t)_i^r}{L_i^r} \right|}{\sum_{t=21}^{24} \frac{N(t)_i^u}{L_i^u}} \quad (4S)$$

$$E_i = \max \left\{ \frac{\sum_{t=21}^{24} L_i^u \times N(t_0) \times e^{\alpha(t-t_0)}}{\sum_{t=21}^{24} N(t)_i^u} ; \frac{\sum_{t=21}^{24} N(t)_i^u}{\sum_{t=21}^{24} L_i^u \times N(t_0) \times e^{\alpha(t-t_0)}} \right\} \quad (5S)$$

in which L_i^u is the length of water mains of group i in the entire utility, L_i^r is the length of water mains of group i in region Z, $N(t)_i^u$ is the number of breaks in group i at year t in the entire utility, and $N(t)_i^r$ is the number of breaks in group i at year t in region Z.

The unbiased standard errors S_i of the nonlinear regression account for the variability in the data from each group in region Z. The factors L_i account for the difference in magnitude between the groups. The square root of the length ratio was taken, rather than the ratio itself, because it seemed that the variability in the data was not directly proportional to the length of the water mains. The factors D_i account for possi-

ble systematic differences between region Z and the entire utility. Finally, the factors E_i account for the quality of prediction from region Z to the entire utility.

How the Herz distribution was derived

The heuristics described earlier in this sidebar could not be applied to the Herz distribution because they resulted in parameters outside of the appropriate domain in some of the groups. Thus, the Herz distribution for the various groups was derived by assuming the following.²²

- The median of the distribution is equal to the time of optimal replacement (Table 6).

- The aging factor parameter, which depicts how slowly aging begins, was assumed to be 1,000 (1/year) for all groups. This value indicates a fairly slow deterioration process, which results in a distribution that is approximately symmetrical about the mode, the mean, and the median (Figure 5).

- The resistance time parameter, which depicts the time lag between the time the pipe was installed and the time at which pipe replacement commences, was taken for each vintage so that replacements commence at 1930, 1950, 1960, 1970, and 1990 for vintages A, B, C, D, and E, respectively (compatible with the Gumbel and Weibull assumptions).

work performed on statistical analysis of water main breaks is provided in this article.

Studies show relationship between breakage rate, age

Shamir and Howard⁷ reported an exponential relationship between pipe breakage rate and age. They introduced an economic analysis to determine the optimal pipe replacement time. Walski and Pelliccia⁸ applied the same exponential relationship to a data set that was divided into pit-cast- and spun-cast-iron mains. Subsequently, they introduced correction factors to fine-tune the exponential function. Clark et al⁹ developed a linear equation to determine the time from pipe installation to the first break and an exponential equation to determine the breakage rate after the first break. Both equations were multivariate, comprising such factors as pipe age, type, soil, and water pressure.

O'Day¹⁰ summarized and compared various studies conducted before 1982 on water main breakage. He concluded that pipe condition could be predicted from age, if geographical location was considered as an additional contributing factor. Studies by Clark and Goodrich¹¹ and Habibian¹² support this observation. Kettler and Goulter¹³ reported a strong inverse linear correlation between main breakage rate and pipe diameter and a moderate linear correlation between breakage rate and pipe age. Andreou et al^{14,15} proposed a probabilistic approach consisting of a proportional hazards model to predict failure probabilities of pipes in the early stages of deterioration and a Poisson-type model for the later stages. The independent variables for the model could be case-specific, e.g., pipe diameter and operating pressure. They acknowledged that the base function for the proportional hazard model (the function that relates the probability of breakage to pipe age) could also vary within a distribution system, a phenomenon that would necessitate layering the data set into two or more groups. Walski¹⁶ introduced additional cost factors (e.g., water loss, leak detection, and repair) to essentially the same economic analysis proposed by Shamir and Howard.⁷

Goulter and Kazemi¹⁷ and Goulter et al¹⁸ observed temporal and spatial clustering of water main failures. A nonhomogeneous Poisson distribution model was proposed to predict the probability of subsequent breaks in a pipe (given that the first break had already occurred). Male et al¹⁹ applied a simulation model to an inventory of water mains in New York City to analyze different long-term pipe replacement strategies. The water main inventory was partitioned into

FIGURE 1 Diagram of validation process to establish data groups for water main breakage analysis

bundles of pipes with uniform characteristics, breakage rate growth patterns were derived, and different replacement strategies were compared. Shamir and Howard's analysis⁷ was expanded to include economic analysis on an infinite series of pipe replacement cycles rather than on an arbitrary planning period.²⁰ This approach allows consideration of the replacement pipe maintenance costs and avoids the need to calculate residual values at the end of the analysis period.

Methodology discusses how to assess capital expenditures with limited data

This article describes a methodology to assess the capital expenditures required for a major water utility that had limited historical data on water main breaks. The proposed methodology draws on some of the work described earlier to provide a rational and consistent treatment of scarce data. It is not uncommon for major utilities to have incomplete data because the regions or districts that currently make up the utility jurisdiction may have had different record-keeping practices or may not have kept records at all. It is even possible that the details on the historical data on water mains may have improved with time as it became evident that the data provided valuable information on the performance of the water distribution system.

The first step in this methodology was to group data on water main breaks so that pipe breakage could be predicted with greater statistical significance for relatively homogeneous water main groups. This is an important step because the establishment of financial resources required for future needs depends on the ability of the model to realistically predict future water main breaks. The second step was to establish breakage rate patterns for the various groups. An exponential increase in breakage rate over the years was assumed, and a nonlinear regression pro-

TABLE 1 Data on water main distribution characteristics

Characteristic or Data Field	Data Descriptor or Comment
Segment identification (ID)	Typically corresponding to the street ID
Region name	Name of the region (nine in total)
Segment length	In fact the street length was provided and was assumed to be equal to the corresponding water main segment length.
Construction period	Typically in 10-year ranges (e.g., 1910–19, 1920–29); water mains installed before 1899 and after 1980 were taken as single age groups.
Replacement period	If applicable
Soil type	Five types—clay, sand, silt, rock, and peat
Segment diameter	In inches
Breakage repair	Number of breakage repairs in each of the years 1993–96 in the entire utility and 1973–96 in the data for region Z
Operating pressure	Average values in 69-kPa (10-psi) ranges (e.g., 276–345 kPa [40–50 psi], 345–414 kPa [50–60 psi])
Leaks	Leaks found during the 1995–96 leak detection campaign
Car loading	Number of cars per day in the street segment
Bus loading	Number of buses per day in the street segment
Pavement type	Rigid (concrete) or flexible (asphalt)
Surface condition	Rating (1–5) given to the road surface condition based on visual inspection
Foundation state	Condition of the road foundation—"good" or "poor"

cedure was applied. The third step was to use principles of life-cycle costing to arrive at optimal replacement timing of the water mains in the various groups. The fourth step explored the variation in replacement timing within individual groups. Because data were insufficient to establish true intragroup variation, heuristic methods were used to explore three probabilistic scenarios. The fifth step was to derive investment schedules for these scenarios.

Historical data on main breaks analyzed

Data were available on 16,383 water main segments. The operational activities of the utility are divided into nine regions. The data available for analysis comprised 18,408 water main segments with a total length of 2,406 km (1,487 mi). After the data were screened for incomplete or inconsistent records, there were 16,383 segments, with a total length of 2,038 km (1,260 mi). The relevant data fields available for the water main segments are shown in Table 1. Main break data for eight of the nine regions in the utility were available for only 1993–96; however, one region (referred to as region Z) had main break data for 1973–96, a 24-year period. The data were arranged in two sets: one for the entire utility (including region Z) and one for region Z, comprising 2,430 water main segments with a total length of 298 km (184 mi). Not all data fields were available for every pipe segment, and the available data fields pertain to breakage repairs and not to breaks (there are unknown numbers of occurrences in which a pipe breakage was not repaired or the repair was deferred).

The four years of data for eight of the nine regions were considered insufficient for a statistically significant analysis. Thus the data from region Z were used to derive water main breakage patterns. These patterns were subsequently used to predict future water main breakage for the entire utility. Implicit in this

course of action are two assumptions: (1) that the history of water main breakage repair can be used to predict future breakage, and (2) that the water main breakage repair patterns in region Z may be used to represent the entire utility. The process illustrated in Figure 1 was applied to validate these assumptions.

Analysis was validated on two levels. Two levels of validation were applied in the analysis. In the first level, breakage patterns derived from 20 years (1973–92) of data from region Z were used to predict the number of breaks for the four years 1993–96. These predicted values were then compared with

the actual number of breakage repairs. Having established that 20 years of breakage history could reasonably predict breakage rated in the subsequent four years, breakage patterns were rederived using the entire set of 24 years of data (1973–96) available for region Z. These patterns were then used to predict the number of breaks in the entire utility for the four years 1993–96. The predicted number of breaks was then compared with the actual number of breakage repairs.

Data on daily temperatures for the period 1973–96 for the entire utility were obtained from Environment Canada to observe the relationship, if any, between the breakage rates and freezing indexes. No significant correlation was found between the freezing indexes and the variations in annual breakage rates. Data were not available for a month-by-month analysis of breakage rates relative to freezing indexes.

Data are bundled into homogeneous groups for two reasons. The rationale for dividing the water main breakage data into relatively homogeneous groups of water mains is twofold. First, establishing breakage patterns in homogeneous groups is likely to yield results with a higher degree of confidence. Second, derived breakage patterns were to be mapped from region Z to the entire utility. Thus, mapping of such patterns from a homogeneous group in region Z to its counterpart in the utility is likely to yield a better fit. The bundling of water mains into homogeneous groups has been done by others.^{7,8,19,21}

Establishing homogeneous groups poses two problems. The establishment of homogeneous groups of water mains may pose two problems: (1) determining what characteristics of the water main (Table 1) should be used to form these groups and (2) deciding into how many groups the water mains should be partitioned. Because the groups have to be mutually exclusive, it is important to limit the number of them. In general, the more groups there are, the less data there

are per group, and the statistical analysis becomes less significant. The characteristics that are the most likely to significantly affect water main breakage patterns therefore need to be identified and used as water main grouping criteria.

Grouping characteristics were chosen.

The initial candidate characteristics for grouping criteria were the pipe vintage, soil type, diameter, operating pressure, car loading, bus loading, road type, surface condition, and foundation state. Because the rating of the road surface condition was based on a recent (1994–96) visual survey, it was assumed unlikely to have consistently affected water mains for 24 years. Thus, it was discarded as a grouping criterion. The state of the road foundation was based on tests that had been conducted over several years as the need arose. No information was available about the timing of these tests. Because the road foundation is itself subject to deterioration, these data could not be expected to indicate any consistent effect on water main breakage. Thus, the foundation state rating was also discarded as a potential grouping criterion.

Pipes were classified into vintage ranges. The remaining candidate characteristics, referred to as “variates,” are shown in Table 2. The vintage ranges roughly correspond to the types of water mains in the entire utility. Vintage A includes all the pipes installed before 1930, corresponding largely to pit-cast-iron pipes that had reportedly been installed until the late 1920s. Vintage B includes all the pipes installed between 1930 and 1949. In this period, most water mains were made up of spun-cast-iron pipes with wall thicknesses similar to those of their pit-cast predecessors. Vintage C includes all the pipes installed from 1950 to 1959. In this period of rapidly developing postwar economy, many water mains (more than 600 km [373 mi] in 10 years) were installed. The accelerated expansion was likely to have been accompanied in many cases by lower standards of installation or failure to impose old standards. The pipes installed were mainly constructed of spun cast iron.

Vintage D included all the pipes installed from 1960 to 1979. In the 1960s, the utility started using imported material for backfill instead of the local excavated material used previously. The pipes in vintage D are predominantly made of spun cast iron with a very small proportion of ductile iron. Few water mains were installed during the 1970s, probably because of the energy crisis and the economic slowdown. Finally, vintage E corresponds to all the pipes installed between 1980 and 1996. These pipes were predominantly made of ductile iron.

Analysis of variance (ANOVA) procedure was applied. A two-stage process using a single- and a double-factor ANOVA²² procedure was applied to

TABLE 2 Variates broken down into range or level

Variate	Range or Level
Vintage—year	Up to 1929, 1930–49, 1950–59, 1960–79, 1980–96
Soil—type	Clay, sand, rock, silt, peat
Pipe diameter—mm (in.)	150 (6), 200 (8), 250 (10), 300 (12)
Operating pressure—kPa (psi)	276–345 (40–50), 345–414 (50–60), 414–483 (60–70),
	483–552 (70–80)
Car traffic loading—cars/day	0–2,000, 2,000–4,000, >4,000, unknown
Bus traffic loading—buses/day	0–100, 100–200, >200, unknown

identify the most significant variates. Because the ANOVA procedure assumes that the data analyzed are distributed normally, the raw data for absolute number of breaks could not be used. Instead, the raw data were transformed in the following manner.

As mentioned earlier, one of the fundamental assumptions in the analysis is that the water main breakage rate increases exponentially with age. A nonlinear regression was applied to the entire region Z pipe breakage data set, and the parameters for the exponential function were extracted (this procedure is described in detail later in this article). The deviations, or residuals, of the actual breakage rates from the regression line were assumed to be normally distributed around the regression line and were subsequently used as the input data for the ANOVA procedure. In the first stage, the data were repeatedly partitioned by each of the variates, and a single-factor ANOVA was applied to each bundle scheme. The variates with low *F*-statistics were discarded as potential grouping criteria. Typically, greater variations between the data in various groups yield a greater *F*-statistic. If the *F*-statistic is below a threshold value (which depends on the degrees of freedom in the data and on the level of significance desired), the data in the various groups are said to belong to the same underlying population. Thus, a low *F*-statistic in a given bundle scheme means that this particular variate cannot adequately explain the variation in the data and thus is not a good grouping criterion.

For example, when the data were partitioned by the vintage variate (Table 3), the *F*-statistics of the single factor ANOVA show that the pipe breakage rates differ significantly between the various vintages (better than a 99 percent significance level). The same partitioning scheme and ANOVA procedure were applied to all the variates (Table 4).

The pipe diameter and car and bus loading variates appear to contribute relatively little to the variation in the data. Thus, these variates were dropped as potential grouping criteria (with the exception of 150-mm [6-in.] pipes, as will be discussed later in this article). In addition to the pipe vintage, which is clearly a good candidate, the soil-type variate appears to be significant at the 92.5 percent level. The operating pressure appears to be somewhat less significant (88 percent) as a source of variation. In the second stage, the remaining variates were further

TABLE 3 Breakage rates partitioned into vintage groups and the resulting single-factor ANOVA*

Year	Breakage Rates Per Vintage—breaks/100 km (breaks/100 mi)					
	A	B	C	D	E	
	Length of Main—km (mi)					
	49.5 (30.8)	3.9 (2.4)	165.6 (103.3)	60.5 (37.7)	18.0 (11.2)	
1973	4.04 (6.5)	0 (0)	15.09 (24.2)	13.23 (21.3)	0 (0)	
1974	2.02 (3.2)	0 (0)	13.89 (22.3)	14.89 (23.9)	0 (0)	
1975	2.02 3.2()	0 (0)	16.30 (26.2)	9.92 (15.9)	0 (0)	
1976	8.07 (12.9)	25.58 (41.1)	13.89 (22.3)	16.54 (26.6)	0 (0)	
1977	10.09 (16.2)	0 (0)	15.70 (25.2)	14.89 (23.9)	0 (0)	
1978	16.15 (26)	25.58 (41.1)	13.89 (22.3)	13.23 (21.3)	0 (0)	
1979	18.17 (29.2)	25.58 (41.1)	17.51 (28.1)	18.19 (29.2)	0 (0)	
1980	6.06 (9.7)	0 (0)	12.08 (19.4)	9.92 (15.9)	0 (0)	
1981	4.04 (6.5)	0 (0)	13.89 (22.3)	8.27 (13.3)	0 (0)	
1982	10.09 (16.2)	0 (0)	10.87 (17.5)	16.54 (26.6)	5.55 (8.9)	
1983	6.06 (9.7)	0 (0)	10.26 (16.5)	6.62 (10.6)	5.55 (8.9)	
1984	2.02 (3.2)	0 (0)	13.28 (21.3)	4.96 (7.9)	5.55 (8.9)	
1985	4.04 (6.5)	0 (0)	9.66 (15.5)	6.62 (10.6)	0 (0)	
1986	2.02 (3.2)	0 (0)	9.66 (15.5)	9.92 (15.9)	0 (0)	
1987	4.04 (6.5)	0 (0)	6.04 (9.7)	14.89 (23.9)	5.55 (8.9)	
1988	2.02 (3.2)	0 (0)	19.92 (32)	21.50 (34.6)	0 (0)	
1989	2.02 (3.2)	0 (0)	11.47 (18.4)	19.85 (31.9)	5.55 (8.9)	
1990	2.02 (3.2)	0 (0)	13.28 (21.3)	16.54 (26.6)	0 (0)	
1991	6.06 (9.7)	0 (0)	12.08 (19.4)	28.12 (45.2)	0 (0)	
1992	10.09 (16.2)	0 (0)	19.92 (32)	18.19 (29.2)	0 (0)	
1993	16.15 (26)	0 (0)	18.72 (30.1)	11.58 (18.6)	5.55 (8.9)	
1994	10.09 (16.2)	0 (0)	22.34 (35.9)	11.58 (18.6)	11.10 (17.8)	
1995	4.04 (6.5)	51.15 (82.3)	15.70 (25.2)	24.81 (39.9)	5.55 (8.9)	
1996	18.17 (29.2)	0 (0)	36.83 (59.2)	39.70 (63.9)	38.85 (62.5)	
Count	24	24	24	24	24	
Sum	−97	−139	96	104	−178	
Average	−4	−6	4	4	−7	
Variance	32	175	28	44	48	
Parameter	Variation					
	SS†	df‡	MS§	F	P-Value	F (90 percent)
Between groups	2,961	4	740	11.259	9.61E-08	1.994
Within groups	7,560	115	66			
Total	10,521	119				
*ANOVA—analysis of variance						
†SS—sum of squares						
‡df—degrees of freedom						
§MS—mean square						

reduced by investigating how they interact with one another, using a two-factor ANOVA procedure. The data were partitioned repeatedly into the remaining variates in a pairwise manner; e.g., for the pair “vintage and soil,” the data were partitioned into 25 bundles comprising all combinations of vintages and soil types. Each pair was then analyzed using a two-factor ANOVA (with 24 replications) procedure to ascertain the degree of linearity between pairs of variates and the degree to which their interaction affects the data (Table 5). Data were not available for some of the groups; e.g., for vintage and pressure, there are no pipes in vintage D and vintage E operating at 276–345-kPa (40–50-psi) pressure. Thus, the two-factor ANOVA procedure was applied to the available data only (Table 5).

The following can be observed from Table 5.

- In general, the data were not well balanced among the various levels or ranges of some of the vari-

ates. Thus, when the data were grouped by two variates, some of the groups had no data, whereas others contained few mains with few breaks, resulting in extremely large variances. For instance, one break occurring in a group containing 200 m (656 ft) of mains translates into a breakage rate of 500 breaks/100 km (805 breaks/100 mi) for that particular group. Thus, the ANOVA results should be interpreted with caution, and engineering judgment should be exercised.

- When the data were grouped into vintage and pressure, most of the variation was attributed to the various vintages rather than to the pressure levels or to the interaction between pressure and vintage. These results seem to be reasonable considering that the typical operating pressures are moderate and that there are no pumping stations in region Z. Thus, the pressure variate was omitted as a grouping criterion.

- When the data were partitioned into vintage and soil types, both variates as well as their interactions

appeared to contribute to data variability. Thus, both the vintage and soil variates were retained as grouping criteria.

- The pavement variate seemed to explain a significant amount of the variation in the data when the data were bundled into vintage and pavement or soil and pavement. However, a closer look at the data showed that > 85 percent of water mains were under rigid pavement, whereas < 6 percent were under flexible pavement (information on the remaining 9 percent was unavailable). Even then, more than half (55 percent) of the 6 percent of the mains under flexible pavement were from vintage E, with the rest divided among vintages A, C, and D. This significant imbalance in the data suggested that the pavement type is not likely to be an efficient grouping factor. Thus, pavement type was also omitted as a grouping criterion.

In addition to the vintage and soil variates selected for grouping criteria, there was a significant difference in the breakage rates of vintage C between 150-mm (6-in.) pipes and larger-diameter pipes. This difference was used as an additional grouping criterion. Thus, the two-stage screening using one- and two-factor ANOVA procedures (with the exceptions noted earlier) permitted the establishment of 16 data groups (Table 6). In the interest of creating groups of sufficient size, some of the levels or ranges were consolidated in the grouping scheme (e.g., group 6 and group 16—all soil types). There are no 150-mm (6-in.) pipes in vintage E.

Region Z's breakage patterns had to be identified. Once the data groups had been established, the water main breakage pattern associated with each group had to be identified. The breakage pattern analyses were based on water main structural breaks and did not include leaks discovered by leak detection programs.

In addition, the following assumptions were made.

- The breakage rates were assumed to increase exponentially over time. This exponential relationship to time has been widely used in various forms.⁷⁻⁹ The exponential relationship is expressed as:

$$N(t)_i = N(t_0)_i \times e^{\alpha_i(t-t_0)} \quad (1)$$

in which $N(t)_i$ is the number of break repairs per year in group i at year t , $N(t_0)_i$ equals $N(t)_i$ at a base year t_0 for group i , and α_i equals the coefficient of breakage rate growth in group i (year⁻¹).

- The base year (t_0) chosen was 1973 (except for group 16 as discussed later in this article) because it is the earliest year for which breakage repair rates are available in region Z.

- The water main breakage records for region Z contained only structural breaks except for the years

TABLE 4 Single-factor ANOVA* of all candidate variates

Variate	F-Statistic	P-Value	F (90 percent)
Vintage	11.259	9.6E-08	1.994
Soil	2.185	0.075	1.994
Diameter	0.495	0.686	2.144
Pressure	1.991	0.121	2.144
Car loading	1.144	0.335	2.144
Bus loading	0.773	0.512	2.144
Pavement	11.650	4.37E-05	2.381

*ANOVA—analysis of variance

1973–79 and 1996. A leak detection program was undertaken for the entire utility in 1995. Thus about 40 percent of the recorded breakage repairs reported in region Z in 1996 were actually repairs of leaks discovered the previous year. The exponential model has to be applied to data that are consistent in order for the breakage pattern analysis to be credible. Because a leak detection program typically increases the amount of maintenance and because there was no consistency in the leak detection program, the analysis had to include only structural breaks.

Leakage repair records had to be omitted from the breakage repair data. No specific data were available regarding which water main segment was repaired for leaks and which for structural breakage. A uniform distribution was assumed, and the number of breaks in 1996 was reduced by 40 percent in every data group. A leak detection and repair program was also in effect between 1973 and 1979. For lack of specific information, the same 40 percent reduction in the reported breakage repairs was applied uniformly across all groups in the years 1973–79.

A nonlinear regression method based on a weighted least-sum-of-squares algorithm was used to fit an exponential curve to the cumulative number of breaks during 20 years (1973–92) for every data group from region Z. These exponential curve fits were then extrapolated to predict the number of breaks for each data group in region Z for 1993–96. Figure 2 illustrates these breakage predictions in 8 of the 16 water main groups in region Z. In the weighted nonlinear regression, the breakage repair rates of 1973–79 were assigned half the weight of the other years because of the previously described uncertainty in the 40 percent data reduction resulting from leak detection and repair programs.

The predictions for each group generally improve with the increase in the total length of water mains and the number of cumulative breakage repairs. For instance, the exponential model is able to make a good prediction of cumulative breaks for groups 9, 10, and 14 with 57.3, 76.8, and 24.4 km (35, 47, and 15 mi) of water mains and about 150, 150, and 45 breaks, respectively. Conversely, the exponential model significantly underpredicts cumulative breaks for groups 1, 5, and 11 with 2.8, 4.4, and 4.4 km (2, 3, and 3 mi) of water mains and 1, 3, and 7 breaks,

TABLE 5 Two-factor ANOVA* results for four variates

Variation	SS†	df‡	MSS§	F	P-Value	F (90 percent)	Comments
Pressure	181	1	181	2.891	0.091	2.733	Based on vintages A, C, D, and E, and pressure of 50–60 and 60–70 psi (345–414 and 414–483 kPa)
Vintage	5,114	3	1,705	27.234	1E-14	2.114	
Interaction	379	3	126	2.020	0.113	2.114	
Within	11,517	184	63				
Total	17,191	191					
Pressure	334	2	167	1.012	0.365	2.328	Based on vintages A, C, and D, and pressure 50–60, 60–70, and 70–80 psi (345–414, 414–483, and 483–552 kPa)
Vintage	4,605	2	2,302	13.943	2E-06	2.328	
Interaction	3,744	4	936	5.668	0.0002	1.972	
Within	34,182	207	165				
Total	42,865	215					
Soil	1,045	2	523	4.675	0.01	2.322	Based on vintages A, C, D, and E, and soil clay, sand, and silt
Vintage	7,630	3	2,543	22.751	3E-13	2.104	
Interaction	1,395	6	233	2.080	0.056	1.796	
Within	30,853	276	112				
Total	40,923	287					
Soil	1,515	3	505	2.900	0.035	2.104	Based on vintages A, C, and D, and soil clay, rock, sand, silt
Vintage	2,326	2	1,163	6.679	0.001	2.322	
Interaction	2,574	6	429	2.464	0.024	1.796	
Within	48,052	276	174				
Total	54,467	287					
Pavement	2,586	2	1,293	12.266	8E-06	2.322	Based on vintages A, C, D, and E, and pavement rigid, flexible, unknown
Vintage	4,061	3	1,354	12.842	7E-08	2.104	
Interaction	2,343	6	391	3.705	0.001	1.796	
Within	29,095	276	105				
Total	38,085	287					
Pavement	3,359	2	1,679	11.718	2E-05	2.328	Based on soil clay, sand, silt, peat, and pavement rigid, flexible, unknown
Soil	882	2	441	3.078	0.048	2.328	
Interaction	992	4	248	1.731	0.144	1.972	
Within	29,666	207	143				
Total	34,898	215					

*ANOVA—analysis of variance

†SS—sum of squares

‡df—degrees of freedom

§MS—mean square

respectively. Exceptions were cumulative breaks underpredicted for group 4 with 25 km (15 mi) of water main and 24 breaks, overpredicted for group 13 with 16 km (10 mi) of water mains and 60 breaks, and predicted accurately for group 2 with 2.2 km (1.4 mi) and 5 breaks. Reasonably good predictions were obtained for the rest of the groups (6, 7, 8, 12, and 15). As mentioned earlier, the base year for group 16 was not 1973 because vintage E comprises years 1980–96. Thus, 1980 was used as the base year for this group. Because of the lack of specific data, it was assumed that the installation of the 18 km (11 mi) of water mains in group 16 was uniformly distributed over the entire vintage E period (i.e., an installation rate of just more than 1 km/year [0.62 mi/year]).

A total of 170 breaks was predicted for all groups for the years 1993–96 compared with the 204 recorded breaks, i.e., an underprediction of about 17 percent. To ascertain the benefit of analyzing grouped

data, an exponential breakage repair rate curve was fitted to ungrouped data, i.e., all the breakage repairs during the 20 years in region Z consisting of 297 km (184 mi) of mains. The resulting curve predicted 158 breakage repairs for the years 1993–96, which is an underprediction of about 23 percent. Thus, there is merit in conducting the analysis by groups to predict future breakage rates. The comparative analysis between grouped and ungrouped data is illustrated in Figure 3.

Breakage rates were predicted for the entire utility. The next step in the analysis was to attempt to use the historical data from region Z to predict breakage rates for the entire utility. The exponential breakage repair rate curves were rederived for all 16 groups using the data from the entire period of 24 years of breaks in that region. The first seven years of data (1973–79) were again assigned half the weight using the nonlinear regression analysis. The data for

FIGURE 2 Breakage patterns for groups 1–8 (derived from 20 years of data) in region Z

the most recent four years (1993–96) were assigned double the weight because data collection procedures have improved and thus there is more confidence in the integrity of the data. The resulting curves were normalized for the water main lengths in the respective groups and were then applied to the corresponding groups in the entire utility, including data from region Z. Table 6 provides the corresponding parameters as well as breakage rate prediction for 1997–2000. Actual breakage data for 1997 and 1998 were not available for comparison. Table 6 and Figure 4 illustrate the comparison between predicted and actual number of breaks in the four-year period between 1993 and 1996.

Although the total prediction for the number of break repairs in the entire utility in the four-year period is close to the actual number, the accuracy of prediction is not consistent among the groups. In general, the prediction is better for a given group when the ratio between the length of pipe for the entire utility and the length of pipe for the group in region Z is < 7.5 . Groups 1 (a good prediction with a length ratio of 54) and 15 (a poor prediction with a length ratio of 2) are exceptions.

Optimal replacement timing results in minimal life-cycle costs

The optimal replacement timing of a water main is that which results in minimal life-cycle costs, sub-

FIGURE 3 Comparison of breakage predictions in region Z between grouped and ungrouped (aggregate) data

used in this article is described in detail in the literature.²⁰ Each group was initially analyzed as though it were a single, uniform water main (variations within groups are addressed later).

Assumptions used to calculate optimal replacement timing. The following assumptions were made to calculate the optimal replacement timing of water mains in each group.

- The utility had reported little deterioration in hydraulic capacity from corrosion or tuberculation in the inner surfaces of most water mains. It was thus assumed that for this particular utility, nonstructural rehabilitation (internal lining) was not likely to be a viable rehabilitation alternative. Thus, the components considered for life-cycle cost were breakage repair and pipe replacement only.

- The parameters $N(t_0)_i$ derived from the 24 years of data (17 years in group 16) and normalized for length were assumed to be the 1973 breakage rates of the respective groups (except for group 16 as discussed earlier).

- The parameters α_i derived from the 24 years of data (17 years in group 16) were assumed to be constant for the respective groups. A negative value (Table 6) for the parameter α_i was assumed to be incorrect because of insufficient data and because a negative α_i value indicates

a decrease in the breakage rate and implies the pipe will last forever. The economic analysis was conducted with a minimum value of 0.001 for α_i , which corresponds to a slowly increasing (nearly constant) breakage rate.

- The average cost of replacing existing pipe was determined by the utility to be \$442/m (\$134/ft), including renewal of service connections for all pipes 150–300 mm (6–12 in.) in diameter. This figure was used for all the groups except group 6 (composed strictly of 150-mm [6-in.] pipes), which was taken as \$340/m (\$103/ft) because it was assumed that these water mains would be replaced by pipes no greater than 200 mm (8 in.) in diameter. The pipe breakage cost was estimated to be \$5,780/event, including costs

ject to some set performance criteria. The life-cycle cost of an existing water main consists of maintenance and rehabilitation–renewal costs. Maintenance costs comprise primarily breakage repair costs, in which external costs such as water loss, damage repair, and social costs could be incorporated.¹⁶ As main replacement is delayed, the total (discounted) cost of maintenance increases because of deterioration, whereas the cost of replacement decreases because of discounting. The total cost typically forms a convex curve with a point of minimum cost. The rehabilitation–renewal timing of this minimum-cost point is referred to as the optimal replacement timing. The optimal replacement timing of a water main can be determined in several ways,^{7,8,16,20} and the method

related to the actual repair and some unspecified indirect costs. The cost of water lost through leaks¹⁶ was not considered because the utility had reported a very low water production cost.

The optimal group replacement timing was calculated using an infinite replacement cycle concept developed by Kleiner.²⁰ This concept considers the costs of the current pipe breakage repair, the replacement pipe, the replacement pipe breakage repair, and the cost of its replacement and so forth in perpetuity. The initial breakage rate parameter $N(t_0)$ and breakage rate growth parameter α for the replacement pipes were assumed to be equal to those of group 16 parameters because group 16 contains ductile-iron pipes—the utility's replacement pipes of choice in the future. The optimal calculated group replacement timing is provided for discount rates of 4 and 6 percent (Table 6).

The initial breakage rate $N(t_0)$ indicates the group breakage rate at the beginning of the analysis period, i.e., 1973 for groups 1–15 and 1980 for group 16. The breakage rate growth coefficient α indicates the average annual increase in the group breakage. In the following discussion, the breakage rate is assumed to be a surrogate measure of water main structural deterioration rate. Thus, these two terms are used interchangeably.

Groups vary in breakage rates and replacement timing. The following observations can be made from this study.

- Group 1, with 251 km (155 mi) of the oldest pipes laid in clay soil, seems to have had a slow rate of deterioration until the early 1970s. This rate has increased significantly in the past 24 years. Although the calculated breakage growth rate parameter was quite large, indicating a fast deterioration rate, these parameters resulted in a fairly good prediction (222 breaks predicted versus 173 breaks observed). Furthermore, the calculated replacement time for this group indicates replacement at a median age of 95 years, which is not unreasonable. However, the high breakage growth rate was obtained from a relatively small group in region Z with a relatively small number of historical breaks, which warrants cautious interpretation.

- Group 5, with 20 km (12 mi) of 45–65-year-old pipes laid in silty soil, seems to display the same characteristics as group 1. Similar to those for group

FIGURE 4 Predicted and actual numbers of breakage repairs, by group, for the entire utility 1993–96

1, the breakage rate parameters have been derived based on a relatively small number of historical breaks.

- Groups 2, 3, and 4, with 112, 277, and 197 km (69, 171, and 122 mi), respectively, of pipes laid before 1929 (vintage A), have remarkably low rates of deterioration for such old cast-iron pipes. However, these specific groups have the highest degree of discrepancy between the predicted and actual number of breaks (Table 6). These large and inconsistent discrepancies could be explained as follows. Vintage A consists of mains installed before 1930. Although no mains were installed before 1900 in region Z, more than 231 km (143 mi) of water mains (about 28 percent of vintage A) in the entire utility were installed in the previous century. This lack of equivalency could be the reason for some of the inconsistencies in the predictions for the vintage A groups.

- Groups 6–10, with 609 km (376 mi) of pipes laid in the 1950s, seem to have a relatively high initial breakage rate with a relatively slow deterioration rate. Although these groups showed relatively high breakage rates, it seems that these rates are increasing very slowly, resulting in relatively late replacement timing.

- Groups 11 and 15, with 63 and 27 km (39 and 17 mi) of pipe laid mainly during the 1960s in clayey soils and peat, respectively, show a relatively high deterioration rate. Pipes of the same vintage within groups 12 and 13—with 25 and 53 km (15 and 33 mi) laid in rock and sandy soils, respectively—show a relatively low deterioration rate. Group 14, with 85

TABLE 6 Group characteristics, derived parameters, predicted and actual breaks, and optimal replacement timings

Group	Vintage†	Soil	Diameter mm (In.)	Length of Water Mains		Parameters Derived From 24 Years of Data		Number of Breaks From Entire Utility 1993–96		Replacement Timing* at Discount Rate		Average Breaks Predicted for Entire Utility 1997–2000	
				Region km (mi)	Utility km (mi)	$N(t_0)$ breaks/ km (mi/year)	α year ⁻¹	Predicted total/ four years	Actual total/ four years	4 percent years	6 percent years	Rate breaks/ km (mi/ year)	Total breaks/ year
1	A	Clay	All	4.4 (2.7)	251 (156)	0.003 (0.005)	0.188	222	173	13	15	0.47 (0.75)	118
2	A	Rock	All	2.2 (1.4)	112 (700)	0.090 (0.144)	0.040	100	31	69	79	0.26 (0.42)	29
3	A	Sand	All	18.0 (11.2)	277 (172)	0.042 (0.067)	0.018	70	140	>100	>100	0.07 (0.11)	19
4	A	Silt	All	25.0 (15.5)	197 (122)	0.074 (0.119)	-0.024†	34	106	>100	>100	0.04 (0.06)	8
5	B	Silt	All	2.8 (1.7)	20 (12)	0.003 (0.005)	0.176	11	10	17	19	0.27 (0.43)	6
6	C	All	150 (6)	5.6 (3.5)	75 (47)	0.076 (0.122)	0.051	72	159	48	56	0.29 (0.46)	22
7	C	Clay	150 (> 6)	11.6 (7.2)	141 (88)	0.087 (0.140)	0.024	83	132	>100	>100	0.16 (0.26)	23
8	C	Rock	150 (> 6)	13.4 (8.3)	94 (58)	0.098 (0.158)	0.027	67	48	>100	>100	0.20 (0.32)	19
9	C	Sand	150 (> 6)	57.3 (35.6)	129 (80)	0.080 (0.129)	0.044	110	87	64	73	0.25 (0.40)	33
10	C	Silt	150 (> 6)	76.8 (47.7)	170 (106)	0.072 (0.116)	0.034	106	110	92	>100	0.18 (0.29)	30
11	D	Clay	All	4.4 (2.7)	63 (39)	0.022 (0.035)	0.104	57	31	25	29	0.34 (0.54)	22
12	D	Rock	All	4.8 (3.0)	25 (16)	0.059 (0.095)	0.036	14	12	91	>100	0.15 (0.24)	4
13	D	Sand	All	16.0 (10)	53 (33)	0.134 (0.216)	0.028	53	40	96	>100	0.28 (0.45)	15
14	D	Silt	All	24.6 (15.3)	85 (53)	0.045 (0.072)	0.055	53	52	55	63	0.20 (0.32)	17
15	D	Peat	All	10.7 (6.6)	27 (17)	0.040 (0.064)	0.096	38	23	23	27	0.51 (0.82)	14
16	E	All	150 (> 6)	18.0 (11.2)	231 (143)	0.009 (0.014)§	0.144	135	84	24	27	0.41 (0.66)	94
Total				295.7 (183.7)	1,952 (1,212)			1,225	1,238			0.18 (0.30)	360

*Optimal replacement timing

†Vintages: A—before 1929, B—1930–49, C—1950–59, D—1960–79, E—1980–96

‡This parameter was adjusted to 0.001 for the calculation of the optimal replacement timing.

§Parameter is adjusted to reflect uniform increase in water main length in years 1980–96.

km (53 mi) of the same vintage D pipes laid in silty soils, shows a medium deterioration rate. This difference in the breakage rate in clayey and silty soils is consistent with external corrosion rates and the performance of water mains that have been observed in other studies.^{10,23} The parameters for groups 11 and 12 were also derived from relatively small numbers of historical breaks.

• Group 16, with 231 km (143 mi) of the newer (vintage E) ductile-iron pipe, shows an initial low breakage rate but a relatively high rate of deterioration, resulting in a recommended short replacement timing. Similar observations pertaining to relatively high breakage rates in unprotected ductile-iron pipes have been reported.^{24–28} However, information on the installation rate of the water mains in this group was not available, and it is likely that the replacement timing is sensitive to the assumed average installation rate. Further, group 16, like group 1, is a large group whose break-

age rate parameters have been derived based on a relatively small number of historical breaks.

• The total cost curve can be quite flat in the vicinity of its minimum value, depending on the breakage rate parameter α , the discount factor, and the difference between them. The calculated optimal replacement times represent a recommended time range rather than a rigid number. This concept is further developed in the following discussion about water main variability within the groups.

Variations also exist within groups

Aside from the apparent variations in the optimal replacement timings between the groups, variations also exist among single pipe segments within any individual group. This is expected in such large groups containing from 20 to 277 km (12 to 171 mi) of water mains. These intragroup variations are addressed in the form of probability distributions.

TABLE 7

Projection of 30-year investment needs for water main replacement—Weibull distribution (by group), and Gumbel and Herz (totals only)

Group	Investment for Years—In millions of dollars						Cumulative Investment for Year—In millions of dollars					
	0-5	5-10	10-15	15-20	20-25	25-30	5	10	15	20	25	30
1	24.2	17.8	23.2	23.1	15.6	6.0	24.4	42.1	65.2	88.4	104.0	110.1
2	6.3	1.1	1.1	1.2	1.2	1.3	6.1	7.4	8.8	10.2	10.8	12.9
3	0.0	0.0	0.06	0.06	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.6
4	0.2	0.1	0.1	0.2	0.2	0.4	0.0	0.0	0.6	0.6	0.6	1.3
5	0.0	0.2	1.9	6.1	0.6	0.0	0.0	0.0	2.0	8.1	8.8	8.8
6	0.06	0.1	0.2	0.4	0.8	1.4	0.0	0.0	0.6	0.6	2.0	3.4
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	0.0	0.0	0.0	0.06	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.6
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	7.5	3.1	3.3	3.2	2.9	2.5	7.4	10.8	14.2	17.0	20.4	22.4
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	0.0	0.0	0.0	0.0	0.06	0.4	0.0	0.0	0.0	0.0	0.0	0.6
15	1.2	2.3	3.9	3.3	0.8	0.06	1.3	3.4	7.4	10.8	12.2	12.2
16	0.06	0.2	1.2	3.7	8.9	17.2	0.0	0.0	1.3	5.4	14.2	31.2
Total—Weibull	39.7	25.1	35.2	41.6	32.0	29.9	39.4	64.6	99.9	142.1	174.0	204.0
Total—Gumbel	42.0	25.0	37.8	45.6	31.2	26.7	42.1	67.3	105.4	150.9	182.2	208.7
Total—Herz	55.3	21.6	30.7	44.2	53.5	47.0	55.0	76.8	107.4	151.6	205.3	252.2

The subject of lifetime probability distribution of water mains has been addressed by Herz²⁹ and subsequently by Deb et al.²⁴ In these two articles, the elusive issue of the end of life of a water main is determined by the past experience of the utility (i.e., a retrospective view of the age distribution at which pipes have been replaced by the utility in the past²⁹) or by the perception of professionals employed by the utility.²⁴ In the analysis presented in this article, the end of life of a water main is assumed to be the end of its economic life. Thus, the optimal replacement time of a main is the end of its life.

In the absence of data, a heuristic procedure was applied to fit three probability distributions to the group data, namely, the type I asymptotic distribution for minima (Gumbel distribution), Weibull, and the Herz distributions. The first is a left-skewed probability distribution depicting a probability of failure, which rises slowly and then diminishes quickly (as the population is consumed). The Weibull distribution has been used extensively in reliability studies and lifetime models. Depending on the values of its parameters, the Weibull distribution can be symmetrical or left- or right-skewed (as will be seen later in this article, the data in all the groups assumed a left-skewed Weibull distribution). The Herz distribution has been applied specifically to water main populations^{24,29} and is essentially a right-skewed probability distribution in which the degree of skewness is subject to parameter selection (in field applications, it was often found to be close to symmetrical²⁵).

The three probability distributions were heuristically applied to each of the 16 groups in the following manner. The median lifetime was assumed to be equal to the respective optimal replacement timing of

each group (Table 6). This assumption means that half the pipes in each group are assumed to have been replaced when the corresponding optimal replacement timing has been reached. The standard deviations of the three distributions for each of the 16 groups were calculated based on the level of uncertainty involved in predicting breakage rates both in region Z and in the entire utility. The derivations of these standard deviations are described in detail in the sidebar on page 48. The distribution of water main replacements of group 1 is illustrated as an example for all three distributions in Figure 5.

For group 1, the optimal replacement timing is 13 years corresponding to a discount rate of 4 percent (Table 6), which translates to the year 2010. The examples in Figure 5 demonstrate the concept of the distribution median being at this optimal replacement timing. It is also shown that water main replacements in group 1 are assumed to commence in year 1930, which is the year subsequent to the last year of vintage A. This initial very low replacement rate accounts for rare occasions in which a relatively young pipe may have been replaced. For vintages B and C, the year of replacement commencement was similarly assumed to be subsequent to the last vintage year. The commencement year of replacement for vintage D was assumed to be 1970 because most water mains in this vintage were installed in the 1960s. Similarly, the year of replacement commencement for vintage E was assumed to be 1990, which is 10 years after the first pipe in this vintage was installed.

Investment needs were projected

The required investment schedules were compiled from the predicted replacement distributions of

FIGURE 5 Example of water main replacement distribution for group 1 at a 4 percent discount rate

all groups. If, for example, the area under the probability distribution curve (Figure 5) of a certain group, e.g., between year 2005 and 2010, depicted a proportion of 10 percent, it was assumed that in this particular group 10 percent of the water mains would have to be replaced in this period. The projected required investments for the next 30 years are depicted in five-year increments (Table 7). A group breakdown is provided for the Weibull distribution. For the other two distributions, only the total investments are provided for comparison. It was assumed that the proportion of water mains that should have been replaced (but were not) according to the schedule before the time of analysis would be replaced in the first five years of the replacement plan.

The Gumbel and the Weibull distributions produce similar investment schedules throughout the 30-year projections. All three distributions produce similar projections in the first 20 years, with the Herz distribution displaying a heavier investment in the first five years. In subsequent years, the Herz distribution projects an increasingly higher investment. The average annual level of investment in the water distribution system projected by both the Herz and Gumbel distributions is about 0.71 percent of the total replacement value of the system in the first 20 years (Weibull is slightly lower at 0.66 percent). This value is higher in the first five years (1.04, 0.79, and 0.75 percent for the Herz, Gumbel, and Weibull distributions, respectively) because of the backlog of replacements that are to be undertaken during that time.

The average annual water main replacement rate is estimated at 0.6 percent in Canada³⁰ and 0.5 percent in the United States.³¹ These values show that at the current replacement rate, the long-term mean age of water mains at replacement is at least 166–200 years, which is very high considering the fact that many utilities are experiencing an infrastructure “crunch” after

less than 100 years since their first water mains were installed. Considering that the utility has had a limited water main replacement program in the past few decades and that the existing 0.6 percent replacement average may be lower than is actually required, the projected replacement needs may err on the low side.

There are two possible reasons for these relatively low figures. The first is that the breakage rate analysis was performed only on breaks that have been repaired. Unrepaired breaks were not reported and thus were not included in the data analyzed. Furthermore, leaks identified through leak detection programs could not be included

in the analysis because of inconsistent leak detection programs over the years. It is very likely that the required investments would be higher if leakage and leakage repair could be consistently considered in the analysis. The second possible reason for the relatively low investment is that water main deterioration in the utility under study is slower than the Canadian average. It was thus suggested that the projected investment figures be treated as a lower bound for the actual required investments in the future. More accurate projections would be possible after additional data are collected and analyzed.

Conclusions

A methodology to estimate distribution system investment needs for an entire utility was developed on the basis of historical water main breakage records available from only one of the nine operational regions within the city. The methodology was developed in five steps: (1) bundle data on water main breaks into homogeneous groups, (2) establish breakage rates for each group and project future breakage rates, (3) use projected breakage rates to determine the economic life of water mains, (4) explore the variations of water main life within the respective groups, and (5) determine the investment required for water main replacement. The methodology described in this article is intended for financial planning and does not provide an operational main renewal plan. Once a specific group of pipes has been identified as requiring attention, a prioritization scheme for the renewal of the pipe segments within the group would be required. Methods such as water audits and the application of nondestructive testing technologies might be used to identify the specific pipe segments. As more data become available, the application of single pipe-based probabilistic analyses as proposed by Andreou¹⁵ may be possible.

References

1. DOLEAC, M.L.; LACKEY, S.L.; & BRATTON, G. Prediction of Time-to-failure for Buried Cast-Iron Pipes. Proc. 1980 AWWA Ann. Conf., Atlanta.
2. BAHMANYER, H.G. & EDIL, T.B. Cold Weather Effects on Underground Pipeline Failures. Proc. Conf. on Pipelines in Adverse Environments, San Diego (1983).
3. KUMAR, A.; MERONKLY, E.; & SEGAN, E. Development of Concepts for Corrosion Assessment and Evaluation of Underground Pipelines. US Army Corps of Engrs., Construction Engrg. Res. Lab., Champaign, Ill. (1984).
4. HABIBIAN, A. Effects of Temperature Changes on Water Main Breaks. *Jour. Transportation Engrg.*—ASCE, 120:2:312 (1994).
5. AHAMAD, M. & MELCHERS, R.E. Reliability of Underground Pipelines Subject to Corrosion. *Jour. Transportation Engrg.*, 120:6:989 (1994).
6. RAJANI, B.; ZHAN, C.; & KURAOKA, S. Pipe-Soil Interaction Analysis of Jointed Water Mains. *Canadian Geotech. Jour.*, 33:3 (1996).
7. SHAMIR, U. & HOWARD, C.D.D. An Analytic Approach to Scheduling Pipe Replacement. *Jour. AWWA*, 71:5:248 (May 1979).
8. WALSKI, T.M. & PELLICCIA, A. Economic Analysis of Water Main Breaks. *Jour. AWWA*, 74:3:140 (Mar. 1982).
9. CLARK, R.M.; STAFFORD, C.L.; & GOODRICH, J.A. Water Distribution Systems: A Spatial and Cost Evaluation. *Jour. Water Resources Planning & Management*—ASCE, 108:5:243 (1982).
10. O'DAY, D.K. Organizing and Analyzing Leak and Break Data for Making Main Replacement Decisions. *Jour. AWWA*, 74:11:589 (Nov. 1982).
11. CLARK, R.M. & GOODRICH, J.A. Developing a Database for Infrastructure Needs. *Jour. AWWA*, 81:7:81 (July 1989).
12. HABIBIAN, A. Developing and Utilizing Databases for Water Main Rehabilitation. *Jour. AWWA*, 84:7:75 (July 1992).
13. KETTLER, A.J. & GOULTER, I.C. An Analysis of Pipe Breakage in Urban Water Distribution Networks. *Canadian Jour. Civil Engrg.*, 12:286 (1985).
14. ANDREOU, S.A.; MARKS, D.H.; & CLARK, R.M. A New Methodology for Modeling Break Failure Patterns in Deteriorating Water Distribution Systems: Theory. *Advances in Water Resources*, 10:3:2 (1987).
15. ANDREOU, S.A.; MARKS, D.H.; & CLARK, R.M. A New Methodology for Modeling Break Failure Patterns in Deteriorating Water Distribution Systems: Applications. *Advances in Water Resources*, 10:3:11 (1987).
16. WALSKI, T.M. Replacement Rules for Water Mains. *Jour. AWWA*, 79:11:33 (Nov. 1987).
17. GOULTER, I.C. & KAZEMI, A. Spatial and Temporal Groupings of Water Main Pipe Breakage in Winnipeg. *Canadian Jour. Civil Engrg.*, 15:1:91 (1988).
18. GOULTER, I.C.; DAVIDSON, J.; & JACOBS, P. Predicting Water-main Breakage Rate. *Jour. Water Resources Planning & Management*—ASCE, 119:4:419 (1993).
19. MALE, J.W.; WALSKI, T.M.; & SLUTSKI, A.H. Analyzing Water Main Replacement Policies. *Jour. Water Resources Planning & Management*—ASCE, 116:3:363 (May-June 1990).
20. KLEINER Y.; ADAMS, B.J.; & ROGERS, J.S. Long-term Planning Methodology for Water Distribution System Rehabilitation. *Water Resources Res.*, 34:8:2039 (Aug. 1998).
21. WALSKI, T.M. & WADE, R. New York Water Supply Infrastructure Study. Tech. Rept. EL-87-9. US Army Corps of Engrs. (July 1987).
22. GUTTMAN, I.; WILKS, S.S.; & HUNTER, J.S. *Introductory Engineering Statistics*. John Wiley & Sons, New York (3rd ed., 1982).
23. ROSSUM, J.R. Prediction of Pitting Rates in Ferrous Metals From Soil Parameters. *Jour. AWWA*, 61:6:305 (June 1969).
24. DEB, A.K. ET AL. *Quantifying Future Rehabilitation and Replacement Needs of Water Mains*. AWWA Res. Fdn., Denver (1998).
25. HERZ, R. Dégradation et Renouvellement des Infrastructures: Un Modèle de Survie par Cohorte. *FLUX*, 23:21 (Jan.-Mar. 1996).
26. Caproco Corrosion Prevention Ltd. Underground Corrosion of Water Pipes in Canadian Cities Case: The City of Calgary. Rept. prepared for CAN-MET, Ottawa, Ont. (1985).
27. DeROSA, P.J. & PARKINSON, R.W. Corrosion of Ductile Iron Pipe. Engrg. TR 241. Water Res. Ctr., Swindon, England (Oct. 1986).
28. LEI, J. Statistical Approach for Describing Lifetimes of Water Mains—Case Trondheim Municipality. SINTEF Civil & Envir. Engrg. Rept. 22F007.28, Trondheim, Norway (Sept. 1997).
29. HERZ, R.K. Aging Process and Rehabilitation Needs of Drinking Water Distribution Networks. *Jour. Water SRT—Aqua*, 45:5:221 (1996).
30. Canadian Water and Wastewater Assn. Municipal Water and Wastewater Infrastructure: Estimated Investment Needs 1997 to 2012. (Apr. 1997).
31. KIRMEYER, G.J.; RICHARDS, W.; & SMITH, C.D. *Assessment of Water Distribution Systems and Associated Research Needs* (90658) AWWA Res. Fdn., Denver (1994).

About the authors: Yehuda Kleiner is a research officer at the Institute for Research in Construction (IRC) of the National Research Council Canada, 1500 Montréal Rd., Bldg. M-20, Ottawa, ON K1A 0R6. He has BSc and MSc degrees from Technion, Israel Institute of Technology in Haifa and a PhD from the University of Toronto in Ontario. He is a member of the American Society of Civil Engineers. His previous work has been published in *Water Resources Research*. Balvant Rajani is a senior research officer at the IRC.