Static and dynamic effects in prioritizing individual water mains for renewal

Kleiner, Y.; Rajani, B. B.

Publisher’s version / Version de l'éditeur:

CCWI2007 SUWM2007 Conference [Proceedings], pp. 1-8, 2007-09-03

Static and dynamic effects in prioritizing individual water mains for renewal

NRCC-49672

Kleiner, Y.; Rajani, B.

A version of this document is published in / Une version de ce document se trouve dans:
CCWI2007 SUWM2007 Conference, Leicester, UK., Sept. 3-5, 2007, pp. 1-8

Static and dynamic effects in prioritizing individual water mains for renewal

Yehuda Kleiner and Balvant Rajani
Institute for Research in Construction, National Research Council, Ottawa, Ontario, K1A 0R6, Canada

ABSTRACT: The physical mechanisms that lead to pipe breakage are often very complex and not well understood, and data on the physical condition of pipes is scarce. Conversely, many water utilities do maintain records of pipe failure (or repair) events, which can be used for the statistical analysis of breakage patterns to discern the deterioration of water mains.

The structural deterioration of water mains and their subsequent failure are complex processes, which are affected by many factors, both static (e.g., pipe material, size, age, soil type) and dynamic (e.g., climate, cathodic protection, pressure zone changes). Several models exist in the literature, which use various statistical methods to analyze breakage patterns of pipe breakage histories. Some of these models were designed to address individual water mains, while others can handle only relatively large groups of pipes, which are presumed to be homogeneous with respect to their deterioration patterns. Dynamic factors can currently be considered only in one a model that was designed to deal with pipe groups. While group deterioration analysis is important for high-level renewal planning, operational considerations require the prioritization of individual pipe for renewal within such groups.

The National Research Council of Canada (NRC), with support from the American Water Works Association Research Foundation (AwwaRF) is undertaking research to develop an approach that will allow the consideration of dynamic factors that influence the breakage patterns of individual water mains. This research is entering its third year and in this paper we provide some interesting interim insights and results.

1 INTRODUCTION

The structural deterioration of water mains and their subsequent failures are complex processes, which are affected by many factors, both static (e.g., pipe material, size, age, soil type) and dynamic (e.g., climate, cathodic protection, pressure zone changes). Limited knowledge and scarcity of data render the physical modeling of all distribution water mains impractical. In contrast, statistically derived empirical models can be useful for small-diameter water mains for which low cost of failure does not justify expensive data acquisition campaigns.

Statistical methods to predict water main breaks use available historical data on past failures to identify pipe breakage patterns. These patterns are then assumed to continue into the future in order to predict the future breakage rate of a water main or its probability of breakage. Statistical methods can be classified broadly into deterministic, probabilistic multi-covariate and probabilistic single-covariate (namely time) models. These models are typically applied to grouped data (Kleiner and Rajani, 2001). Deterministic models predict breakage rates using two or more parameters, based on pipe age and breakage history, e.g., Shamir and Howard (1979), Walski and Pellicia (1982) and Clark *et al*. (1982). Many factors, operational, environmental and factors dependent on pipe material, jointly affect the breakage rate of a water main. The population of water mains has to be partitioned into groups that are appreciably uniform and homogeneous with respect to these factors, in order for two or three parameters to capture a true breakage pattern. The partitioning of data into groups, however, warrants careful attention because two conflicting objectives are involved. On one hand the groups have to be small enough to be uniform but, at the same time, the groups have to be large enough to provide results that are statistically significant. Probabilistic multi-covariate models, such as those based on the proportional hazards statistical model (e.g., Andreou *et al*., 1987; Eisenbeis, 1994) or on non-homogeneous Poisson process (e.g., Constantine and Darroch, 1993; Røstum, 2000), can explicitly and quantitatively consider most of the covariates in the analysis. This ability makes them potentially powerful to predict future breakage rates of water mains. It also reduces the need to pre-partition the data into groups, although often some level of partitioning may still be required. Other types of approaches include accelerated failure time (Røstum, 2000) and models that attempt to fit probability distributions to inter-break time durations in pipes, e.g., Gustafson and Clancy (1999), Mailhot *et al.* (2003) and Dridi *et al.* (2005).

All the models described above can only deal with static covariates. Water utilities have observed that operations such as pressure control, cathodic protection (both systematic (retrofit) and opportunistic (hot spot)) and external environment (water temperature and soil moisture deficit) can have a substantial impact on water main failure patterns. Neglecting to account for these effects can lead to inaccurate conclusions, which result in sub-optimal renewal strategies. Kleiner and Rajani (2000, 2004) proposed a model that considers both static and dynamic influences on the breakage pattern of water mains. It is a deterministic model, which works only with groups of pipes. In this paper we report on interim results of research intended to extend this model to consider individual (rather than groups of) water mains to enable effective prioritization of the renewal of these individual water mains.

The rest of this paper is organised as follows: Section 2 provides an introduction to the analysis of historical breakage patterns of groups of water mains, with the consideration of time-dependant (dynamic) effects. Section 3 describes on-going research on ways to "explain" variations in breakage patterns of individual mains, and Section 4 provides a summary and direction of further research.

2 MODELLING DYNAMIC EFFECTS ON BREAKAGE PATTERNS OF GROUPS OF PIPES

The model proposed by Kleiner and Rajani (2000, 2004), named D-WARP (Distribution Water mAins Renewal Planner), addresses the deterioration rates of a homogeneous group of distribution water mains. It considers time-dependent (dynamic) factors such as temperature (in the form of freezing index), soil moisture (in the form of rainfall deficit), and cathodic protection (CP) strategies, including hot spot CP as well as systematic retrofit CP. Non-time-dependent (or static) factors such as pipe characteristics and soil type are implicitly considered through the formation of homogeneous water main groups. The underlying premise is that a homogeneous group of pipes experiences a steady increase in breakage rate (hereafter referred to as "background ageing rate" or simply "ageing rate"), upon which year-to-year variations occur. Some of these variations can be attributed to time-dependent factors. Once background ageing rates is discerned, it can be used to project future breakage rates. In addition, the impact of operational strategies such as schedules of cathodic protection (both hot spot and retrofit) can be superimposed on this background ageing. Subsequently, the life cycle costs of various scenarios operational strategies can be evaluated and fine-tuned to achieve maximum efficiency in resource allocation.

D-WARP uses a general, multi-covariate exponential model to discern breakage patterns while considering time-dependent factors:

$$N(\underline{x}_t) = N(\underline{\overline{x}}_{t_0})e^{\underline{a}\cdot\underline{x}_t^T} \qquad (1)$$

where $\underline{x}_t$ = vector of time-dependent covariates prevailing at time t, $N(\underline{x}_t)$ = number of breaks resulting from $\underline{x}_t$, $\underline{a}$ = vector of parameters corresponding to the covariates $\underline{x}$; $\underline{x}_{t_0}$= vector of baseline $\underline{x}$ values at year of reference t_0. Time-dependent covariates (or "explanatory variables") can be pipe age, temperature, soil moisture, number of effective CP anodes, etc. Parameters $N\left(\underline{\overline{x}}_{t_0}\right)$ and $\underline{a}$ can be found by least square regression or by using the maximum likelihood (ML) method.

The multi-covariate model is applied to groups of water mains that are assumed homogeneous with respect to their deterioration rates. Grouping of water mains is typically based on some or all of the static factors (e.g., by material type, diameter, vintage, geographical location, etc.) for which data are available. Although the mathematical model is not restricted in the number of covariates it can consider, data availability (or rather unavailability) is usually the limiting factor. Currently the only time-dependent covariates that are considered are age, temperature, winter soil moisture, annual soil moisture, hotspot cathodic protection and retrofit cathodic protection. Freezing index is used as a surrogate measure for annual temperature. Two schemes of rain deficit are used as surrogate measures for winter soil moisture and annual soil moisture. Details on these surrogate measures are provided in Kleiner and Rajani (2002).

Hot spot cathodic protection is the practice of opportunistically installing a protective (sacrificial) anode at the location of a pipe repair. These anodes are typically installed without any monitoring and stay in the ground until total depletion, usually without replacement.

Retrofit cathodic protection refers to the practice of systematically protecting existing pipes with galvanic cathodic protection, whereby an anode is attached to each pipe segment (typically 6 m or 20' length), or if the water main is electrically continuous a bank of anodes in a single anode bed is sufficient. Kleiner and Rajani (2004) described in detail how cathodic protection is considered in the multi-covariate model (Equation 1). Figure 1 provides an illustration of an example group of pipes for which a HS CP program was started in 1984.

The break history analysis in D-WARP provides the coefficients a for background ageing, climatic covariates and cathodic protection effects. Once background ageing has been discerned, it is assumed that the group of pipes will continue to age at the same rate. It is also assumed that the cathodic protection effects observed in the past will continue to prevail if CP is continued. Based on these assumptions a forecast of future breakage rates can be made. Note that climate effects are usually not considered in this forecast because they require a credible climate forecast, including temperatures and precipitation

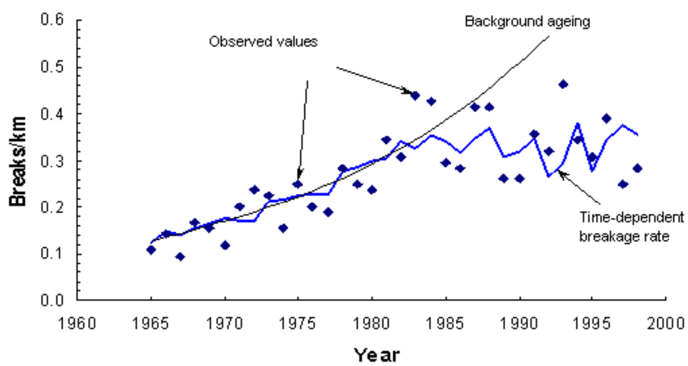

Figure 1. Breakage pattern of cathodically protected pipes (HS CP started 1984)

3 INVESTIGATING VARIABILITIES IN BREAKAGE COUNT OF INDIVIDUAL PIPES WITHIN A HOMOGENEOUS GROUP

The ultimate goal of the research on which this paper reports is to develop an operational tool for network owners and planners to be able to prioritise individual water mains for renewal, while considering both static and dynamic effects that influence pipe deterioration rates. The underlying assumption of our approach is that within a homogeneous group of pipes, some of the variations in breakage rates among individual water mains are a result of irreducible random natural variation (aleatory uncertainty), while some of the variation (epistemic uncertainty) can be explained by the existence of some factors. These factors need to be identified and subsequently expressed as covariates (or "explanatory variables") in a mathematical model.

3.1 *Candidate covariate: pipe length*

In practically all reported analyses of breakage frequency in pipe groups (including those cited in Section 1, as well as others), the aggregate length of the pipes in a group has been used as a normalizing factor. This has the implication that breaks are distributed uniformly along the pipes, which carries the expectation that the number of breaks be directly proportional to the length of pipes. The literature also reflects that pipe length has frequently been used as a covariate capable of "explaining" at least some of the variability observed in individual water mains. Researchers (e.g., Andreou *et al*, 1987; Eisenbeis, 1994; Røstum, 2000, and others) used the log transform or *n*-th root of pipe length as a covariate in their proportional hazards models. These researchers reported various results with respect to the "quality" of pipe length as a covariate. In some water distribution networks, length was found to be statistically significant, while in others it was not. Additionally, pipe length was found to be significant for some pipe materials (CI, PVC) when the number of previous breaks was between 1 and 3, while insignificant when the number of previous breaks was 0 or equal or greater than 4. In other materials (AC) it was found to be not significant at any level of previous breaks. In some cases, when the length was found to be significant, the coefficient (in the exponent) was obtained with values around 0.5. This implied that the hazard was proportional to the square root of the pipe length. Giustolisi *et al.* (2005) used Evolutionary Polynomial Regression (EPR) to find polynomial expressions that predict the breakage rate of pipes. In several data sets length was found to be a good predictor.

These inconclusive results on using length as covariate, motivated us to re-examine pipe length as a candidate covariate for explaining differences in breakage intensity of individual pipes in a homogeneous group. The length of individual water mains is typically in the order of magnitude of no more than several hundreds of meters. A protocol was therefore developed to artificially generate pipes with a wider range of lengths:

a) Extract a homogeneous group of pipes (same diameter, material, vintage, cathodic protection). Total number of individual pipes in the group is N. For each pipe i (i = 1, 2, ..., N), record its length, l_i, and the number of breaks, b_i, observed during the analysis period.
b) Compute the correlation (Pearson) coefficient between l_i. and b_i (individual pipes).
c) Set a constant, say, *MaxPipes* = 2.
d) Generate a positive random integer *NoPipe* where *NoPipe* ≤ *MaxPipes*
e) Generate *NoPipe* positive unique random numbers, say k_1, k_2, …, k_{NoPipe}. each of which ≤ N.
f) Record the sum of lengths $L_j = \Sigma(l_{k_1}, l_{k_2} \ldots)$ and the sum the breaks, $B_j = \Sigma(b_{k_1}, b_{k_2} \ldots)$.

g) Repeat steps (d) to (f) thousand (1000) times. We now have 1000 pairs L_j and B_j (j = 1, 2, …, 1000).
h) Compute the (Pearson) correlation coefficient between L_j and B_j.
i) Repeat steps (d) to (h) for *MaxPipes* = 2, 5, 10, 20, 30, 50, 100, 500 and 1000.

The protocol was applied, with similar results, to three different data sets from three major Canadian cities. Figure 2 illustrates detailed results for one of them, while Figure 3 compares the overall correlations for all three data sets.

Figure 2 demonstrates that while in the macro level (long pipe lengths) there is clear linear correlation between pipe length and number of breaks, in the micro level (shorter pipe lengths) this correlation is completely overtaken by “noise”, which is the natural variation between pipes. In fact, viewing

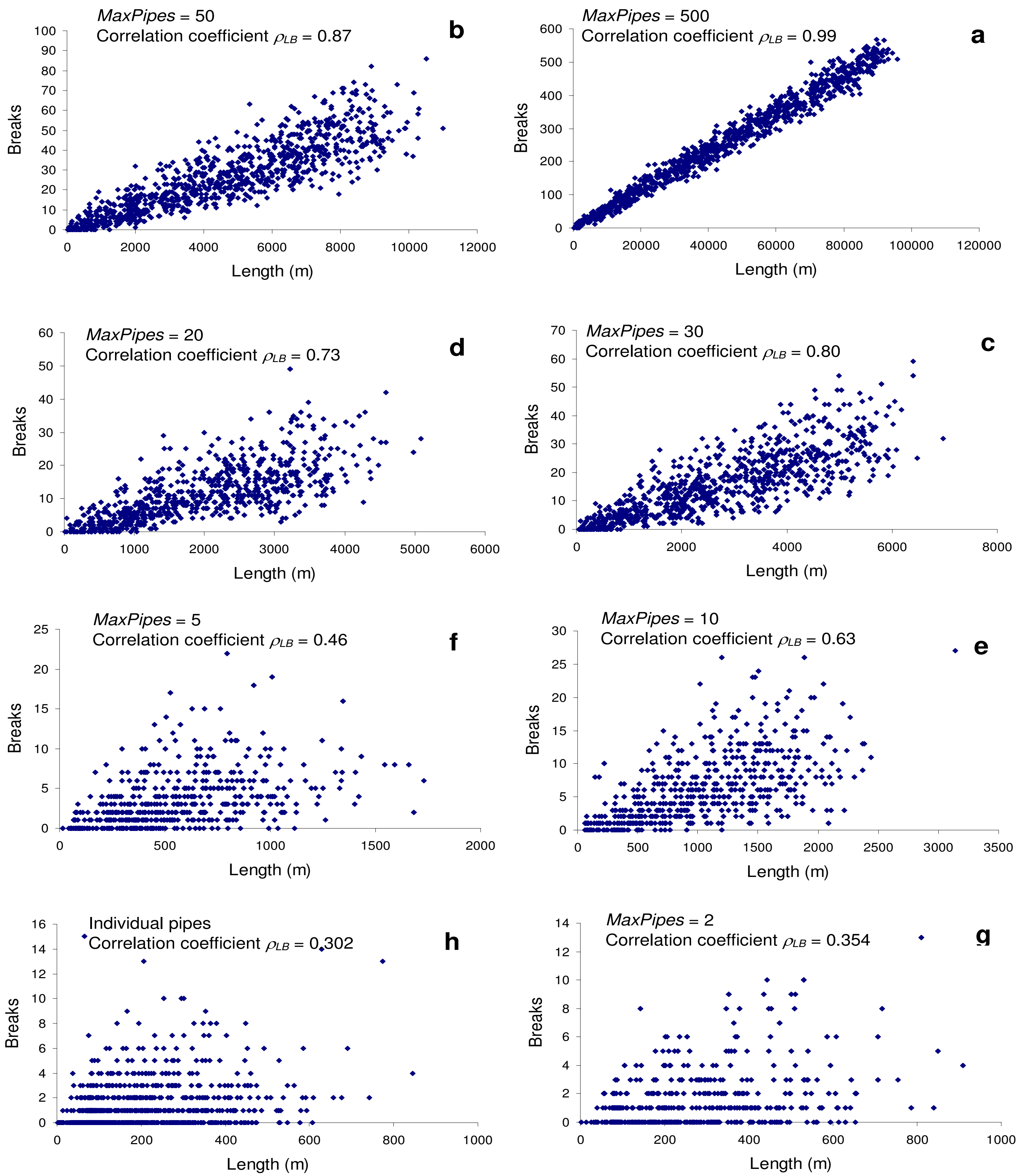

Figure 2. Linear correlation between pipe lengths and breaks: Ottawa dataset CI, 150 mm pipes (ρ_{LB} = correlation length/breaks)

graph a to graph h (top right to bottom left in Figure 2) is akin to zooming in on the bottom left corner of the graph a.

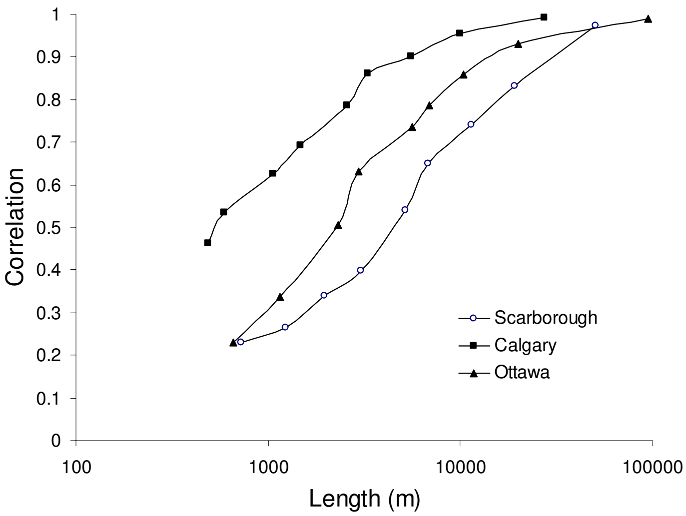

Figure 3. Variation of linear correlation coefficients with pipe lengths in 3 cities

It appears that the intuitive understanding that higher exposure leads to more observed breaks must be true since pipe length is a surrogate for exposure. However, the natural randomness inherent in the relationship between length and breaks is relatively high. Further, while pipe length is a continuous physical property, pipe break is a discrete entity. Individual pipes, whose length might typically vary between a few tens and a few hundreds of meters, typically do not experience too many breaks before they are replaced. This discrete nature of breakage data amplifies the natural randomness in relatively short pipes. Consequently, the randomness or the "noise" in the data all but completely overwhelms any mathematical relationship when comparing individual mains. However, when the variability in pipe lengths is big, the aggregate number of breaks becomes continuous-like in its behaviour and the natural randomness produces "noise" that is much smaller relative to the mathematical relationship and therefore no longer overwhelms this relationship. It should be noted that repeating the same exercise with powers (between zero and unity) and log of lengths yielded similar results.

While Pearson's linear correlation was rather low between length and breaks of individual mains, Spearman's rank correlation proved significantly higher. Spearman's rank correlation values for the individual water mains in Ottawa, Scarborough and Calgary were 0.38, 0.40 and 0.56 compared to Pearson's linear correlation values of 0.23, 0.23 and 0.46, respectively.

3.2 *Candidate covariate: pipe failure history*

Pipe failure history can be characterised by the number of previous failures (*NOPF*), as well as the temporal pattern of these failures, i.e., recency of these failures. Models reported in the literature (e.g., Andreou *et al*, 1987; Eisenbeis, 1994; Røstum, 2000) only considered *NOPF* as a covariate pertaining to pipe failure history. .

To investigate how *NOPF* can be a predictor for future breaks,, we established a reference year and defined time windows for past and future pipe breaks, *WinPast* and *WinFuture*, respectively as as illustrated in Figure 4. The number of recorded breaks in *WinPast* was correlated to the number of recorded breaks in *WinFuture*. This was done for different reference years as well as for variouse t, values of *WinPast* and *WinFuture*. Observed trends (some intuitively expected, others not as much) were:

- Correlations between total number of past breaks and total number of future breaks increases as the length of the *WinPast* increased.
- Correlations between total number of past breaks and total number of future break increases as the length of *WinFuture* increased.
- Correlations between total number of past breaks and total number of future break generally vary for different *RefYears*. However, these variations depend on the lengths of *WinPast* and *WinFuture*. As *WinPast* and *WinFuture* increase, these variations decrease and these variations all but disappear for *WinPast* ≥ 10 years and *WinFuture* ≥ 10.
- The correlation between total number of past breaks and total number of future when *WinFuture* = 1 year is very small (approximately between 0 and 0.2) and varies highly with *RefYear*, regardless of length of *WinPast*. This suggests

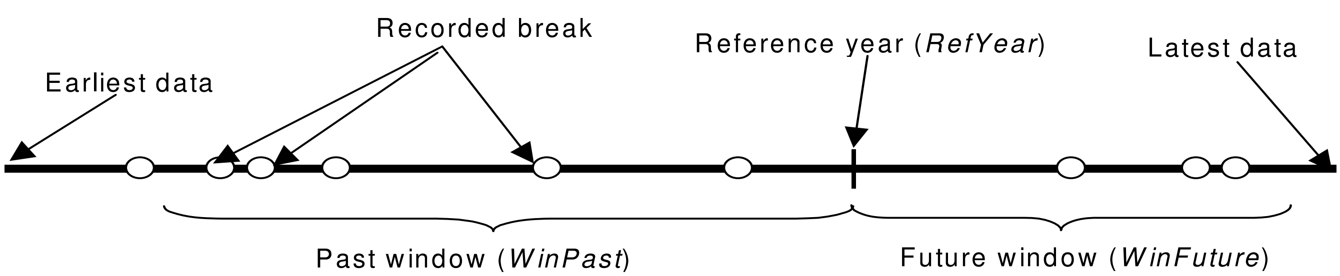

Figure 4. Timeline to explore pipe failure history

that the possibility of predicting next year's breakage in individual water mains, using their past number of breaks is doubtful.

- Correlations between total number of past breaks and total number of future break when *WinFuture* ≥ 5 years and *WinPast* ≥ 15 years have values approaching 0.4.

3.3 *Candidate covariate: geographical clustering*

Water utilities often lack data that are (directly or indirectly) geographically related, such as soil data, overburden characteristics (land development, traffic loading), historical installation practices, groundwater fluctuations, transient pressures, poor bedding, etc. These data, if available, may sometimes help to "explain" variations in breakage rates. Under the hypothesis that geographical clustering of historical breaks could be a surrogate for these often missing data, we examined the viability of using geographical clustering of historical breaks as a possible predictor to explain variability of breakage rates among individual water mains in a homogeneous group of pipes. The examination comprised the following steps:

a) Extract a homogeneous group of N pipes (same diameter, material, vintage, cathodic protection). For each pipe i (i = 1, 2, ..., N) record the number of breaks b^t_i, observed at each year $\boldsymbol{t}$ in the analysis period T (t = 1, 2, …T).
b) For a reference year *RefYear* ($\subseteq T$), select a window of historical years *WinPast* and a window of future years *WinFuture*. *WinPast* comprises T_p years, where the first year is equal to (*RefYear* - T_p + *1*) and the last year equals *RefYear*. *WinFuture* comprises T_f years, where the first year is equal to (*RefYear* + *1*) and the last year equals (*RefYear* + T_f).
c) Based on *WinPast* breakage data create C geographical clusters. A break belongs to cluster c_j (j = 1, 2, …, C) if its Euclidean distance from the centroid of c_j is smaller than its distance from the centroid of every other cluster c_j ($i \neq j$). The clustering algorithm K-Means as described by MacQueen (1967) was used.
d) Partition all pipes into the C clusters, based on their distances from the centroids of the clusters.
e) Create C clustering covariates for each pipe, based on the pipe distance from the centroids of the respective clusters. These covariates are supposed to predict the number of breaks observed during *WinPast*.
f) Train the model on the *WinPast* data by finding C coefficients that minimise the sum of square differences between observed and predicted number of breaks in *WinPast*.

Use the C coefficients to predict breaks for *WinFuture*.

Figure 5 illustrates an example of break clustering (6 clusters) in a group of cast iron 150 and 200 mm water mains, installed in Edmonton, Canada between 1902 and 1945.

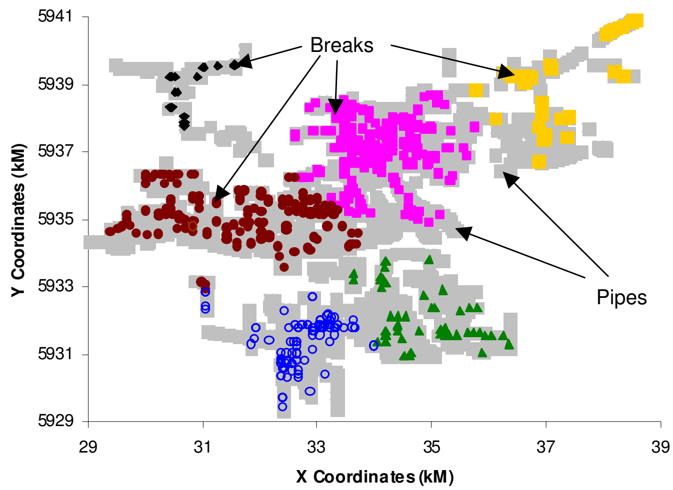

Figure 5. Example of historical breaks partitioned into six clusters

4 COMBINING EXPLANATORY VARIABLE INTO A PREDICTIVE MODEL

Based on the observations reported in the previous section, we concluded that with the given type of data it was more tenable to forecast the relative expected breakage frequency of individual pipes within a homogeneous group of pipes. We therefore set out to address the following challenge: "In a homogeneous group comprising N individual pipes, with available breakage history of T years, find which n pipes are expected to have the highest number of breaks in the next y years". The following procedure is proposed:

a) Select *RefYear* and partition the observation period T into *WinPast* and *WinFuture*. Identify n pipes with the highest number of breaks in period *WinFuture*. Record these pipes in a list named L_f.
b) Similar to step (a) above, identify n pipes with the highest number of break in period *WinPast* and record them in a list named L_1.
c) Identify n pipes with the highest length and record them in a list named L_2.
d) Identify n pipes that the cluster analysis predicted to have the highest number of breaks and record them in a list named L_3.

e) Assign weights W_1, W_2, W_3 to lists named L_1, L_2, L_3, respectively. These weights will serve as initial values to be optimized later.
f) Identify pipes that appear in lists L_1, and L_2 and L_3 and record these pipes in a list named $L_{1,2,3}$. Similarly, identify pipes that appear in lists L_1, and L_2 and record these pipes in a list named $L_{1,2}$. List named $L_{1,3}$ will comprise pipes that appear in lists L_1 and L_3., and list named $L_{2,3}$ will comprise pipes that appear in lists L_2 and L_3.
g) If a pipe appears in a higher order list, remove it from all the corresponding lower order lists. For example, if a pipe appears in list $L_{1,2,3}$ it should be removed from lists $L_{1,2}$, $L_{1,3}$, $L_{2,3}$, L_1, L_2, and L_3. In this way every pipe will appear only once in the 7 lists.
h) Pipes in list $L_{1,2,3}$ are assigned weight $W_{1,2,3} = W_1 + W_2 + W_3$. Pipes in list $L_{1,2}$ are assigned weight $W_{1,2} = W_1 + W_2$, and so on.
i) Rank all pipes (from the 7 lists) by their assigned weight in descending order. The n pipes with the highest weights are those predicted by the model to have the highest number of breaks in the future window. Record these pipes in a list L_W .
j) Denote by H the number of pipes ("hits") that appear in both lists L_w and L_f.
k) Find a set of weights (W_i ; $i = 1, 2, 3$) that maximizes H.

This procedure was applied to several datasets and demonstrated relatively good results. For example, on a set of 1349 150 mm CI pipes (Scarborough), installed between 1945 and 1960, the number of "hits" was $H = 21$ pipes on attempting to predict $n = 75$ pipes with the highest number of breaks.

4.1 *Evaluation of prediction quality*

To evaluate the quality of prediction it is proposed to compare it to the probability of obtaining the same results with a random draw, i.e., what is the probability of randomly selecting n out of N pipes, so that H of the selected pipes ($H \leq n$) truly belong to the n pipes with the highest number of breaks. In our example, $N = 1349$, $n = 75$, $H = 21$, Figure 6 illustrates that obtaining more than 21 "hits" at random has virtually zero probability (1 – cumulative probability).

5 SUMMARY AND CONCLUSIONS

The analysis of historical breakage patterns in a relatively homogeneous group of pipes can provide insight into the expected future trends of the group. However, it is rarely feasible to replace an entire group of pipes due to budgetary constraints, therefore there is a need to prioritise the replacement of individual pipes within such a group. The first step towards prioritisation is to discern differences between the individual pipes, specifically, how to predict that one pipe will fail more frequently than another in the same group?

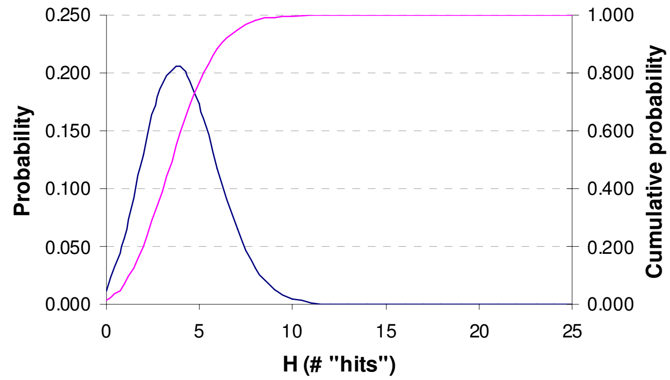

Figure 6. Probability of obtaining results by random selection

In this paper we have reported on the examination of three candidate indicators (covariates), pipe length, number of previous observed failures and pipe and break clustering. A weighted ranking – based procedure was presented, whereby these indicators are used to forecast the pipes with the highest number of breaks in a homogeneous group of pipes. The quality of the prediction was ascertained by comparing the prediction to random selection. More research is planned to test the procedure for consistency across varying time windows and reference years as well as across different data sets.

6 ACKNOWLEDGMENTS

This paper presents interim results of a research project, which is co-sponsored by the AwwaRF, the NRC and water utilities from the United States and Canada. The authors wish to acknowledge the competent programming help of Mr. Jeff Lalonde and the good ideas of Dr. Rehan Sadiq, both at the NRC who work in our group.

REFERENCES

Andreou, S. A., Marks, D. H., and Clark, R. M. (1987). "A new methodology for modeling break failure patterns in deteriorating water distribution systems: Theory." *Advance in Water Resources*, 10, 2-10.
Clark, R. M., Stafford, C. L., and Goodrich, J. A. (1982). "Water distribution systems: A spatial and cost evaluation", *Journal of Water Resources Planning and Management Division, ASCE*, 108, (3), 243-256.
Constantine, A. G., and Darroch, J. N. (1993). "Pipeline reliability: stochastic models in engineering technology and management." S. Osaki, D.N.P. Murthy, eds., *World Scientific Publishing Co.*
Dridi, L,. Mailhot, A., Parizeau, M., and Villeneuve J.P. (2005) "A strategy for optimal replacement of water pipes integrating structural and hydraulic indicators based on a statistical

water pipe break model". *Proceedings of the 8th International Conference on Computing and Control for the Water Industry,* U. of Exeter, UK, September, 65-70.

Eisenbeis, P. (1994). "Modélisation statistique de la prévision des défaillances sur les conduites d'eau potable, PhD thesis, University Louis Pasteur of Strasbourg, collection Etudes Cemagref n° 17, 1994.

Giustolisi, O., Laucelli, D, and Savic D. A., (2005). "A decision support framework fro short time planning of rehabilitation", Proceedings of Computer and Control in Water Industry (CCWI), (1), 39-44.

Gustafson, J-M., and Clancy, D. V. (1999). "Modeling the occurrence of breaks in cast iron water mains using methods of survival analysis" *Proc. AWWA ACE, Chicago*.

Kleiner, Y. and Rajani, B. (2000). "Considering time-dependent factors in the statistical prediction of water main breaks". *AWWA Infrastructure Conference Proceedings*, March 12-15, 2000, Baltimore, Maryland.

Kleiner, Y. and Rajani, B. (2002). "Forecasting Variations and Trends in Water Main Breaks", *Journal of Infrastructure Systems*, *ASCE*, 8, (4), 122-131.

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

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

MacQueen, J. B. (1967). "Some methods for classification and analysis of multivariate observations", *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press*, 1:281-297.

Mailhot, A., A. Paulin, and Villeneuve J-P, (2003). "Optimal replacement of water pipes", *Water Resources Research*, 39, (5), 1136.

Røstum, J. (2000). "Statistical modelling of pipe failures in water networks". PhD thesis, Norwegian University of Science and Technology, Trondheim, Norway.

Shamir, U. and Howard, C.D.D. (1979). "An analytic approach to scheduling pipe replacement." *J. AWWA*, 71, (5), 248-258.

Walski, T. M., and Pelliccia, A. (1982). "Economic analysis of water main breaks." J. AWWA, 74, (3), 140-147.