

## Research Article

# Water Distribution System Pipe Break Failure Modeling and Analysis: Case Study of Kuwait

Jasem M. Alhumoud  and Nourah Almashan

Civil Engineering Department, College of Engineering and Petroleum, Kuwait University, P.O. Box 3569, Safat, Kuwait

Correspondence should be addressed to Jasem M. Alhumoud; jalhumoud@yahoo.com

Received 2 July 2019; Revised 4 February 2020; Accepted 2 May 2020; Published 23 June 2020

Academic Editor: Carlo Rainieri

Copyright © 2020 Jasem M. Alhumoud and Nourah Almashan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Ministry of Electricity and Water (MEW), as well as other water authorities all over the world, is facing a difficult challenge in assessing the physical condition of its distribution systems. Since the majority of the mains are buried, the MEW must rely upon indirect methods, including analysis of repair records. A case study on Kuwait's water distribution system using the techniques of survival analysis is analyzed and investigated for modeling the pipe break failures. The proportional hazard model has the advantage of being able to separate the effects of component deterioration on failure due to aging from the effects of site-specific causes. Another desirable feature is its ability to analyze censored data. The sensitivity of the model parameters to sample size and percent censoring is examined through random sampling of the database. In addition, the proportional hazard model is suitable for describing failure rates of components.

## 1. Introduction

**1.1. General.** In the competition for scarce resources to repair and maintain our public work infrastructure, the use of risk and reliability analysis has been proposed for prioritizing maintenance and repair of constructed systems. One of the hindrances to the application of the models that have evolved to real-life systems is the lack of adequate information on the failure rates of the components that comprise the system.

In the case of water distribution systems, plenty of information is available on failed components. File cabinets at hundreds of water utilities are bulging with records of failed pipes, joints, valves, and pumps. Unfortunately, this reservoir of information has not been efficiently tapped to estimate parameters that are useful for maintenance and repair decisions.

Partly, this situation may be attributed to (1) the fact that relevant data are reported only on the components that have failed, and information on those that are still functioning is neglected; (2) the effect of component aging on failure is not separated from that caused by other explanatory variables.

By neglecting valuable information on unfailed components, a bias is introduced in the modeling of failure.

Information on unfailed components may be considered by using the concepts of censored data analysis. In the second case, because the effect of aging, which is obviously time dependent, is lumped with other causes of pipe failure, which are usually location dependent (i.e., site specific), the failure models obtained have not been too successful. However, some newer models have shown some success in their applications [1–5]. The proposed model overcomes both drawbacks by using the proportional hazard model, originally proposed by Cox and other authors in [6–11]; this model has been used by Andreou et al. [12], Shamsi [13], and also in [14–16].

**1.2. Background of Kuwait's Water Distribution Networks.** Kuwait is located at the north western shore of the Arabian Gulf. This location is categorised as an arid region, which is marked by very scarce rainfall, high temperatures and evaporation rates, and a lack of perennial surface waters. It covers an area of 18,000 km<sup>2</sup>. The average annual rainfall ranges from 40 to 240 mm, and the total annual evaporation rate ranges from 2,500 mm in the coastal areas to more than 4,500 mm inland. Kuwait relies mainly on desalinated water

from the Gulf for potable water. Currently, there are ten desalination plants in Kuwait with a total annual production of  $159 \times 10^9$  imperial gallons (IG) ( $723 \times 10^6 \text{ m}^3$ ). In addition to distilled freshwater, Kuwait has a large supply of brackish groundwater. The Ministry of Electricity and Water (MEW) in Kuwait distributes the brackish water to the consumers through a separate network parallel to the freshwater network. This brackish groundwater is mainly used for irrigation, domestic purposes, and blending with distilled water [17, 18]. Fresh and brackish water are stored in reservoirs and elevated towers. For freshwater, the storage capacities are as follows: (a)  $2.2 \times 10^9$  IG ( $10 \times 10^6 \text{ m}^3$ ) for reservoirs operated by gravity, (b)  $2.1 \times 10^9$  IG ( $9.55 \times 10^6 \text{ m}^3$ ) for reservoirs operated by pumps, and (c)  $55.2 \times 10^6$  IG ( $250,944 \text{ m}^3$ ) for elevated towers. However, for brackish water, the storage capacities are as follows: (a)  $498 \times 10^6$  IG ( $2.2 \times 10^6 \text{ m}^3$ ) for reservoirs operated by gravity, (b)  $41 \times 10^6$  IG ( $186,400 \text{ m}^3$ ) for reservoirs operated by pumps, and (c)  $10 \times 10^6$  IG ( $45,460 \text{ m}^3$ ) for elevated towers. All ground reservoirs are operated by gravity and are built on elevated ground. In addition, freshwater reservoirs are sterilized continuously.

The water distribution network pipes are mainly made of ductile iron with few steel-coated pipes. The diameter of the pipes is 3/4, 1, and 2 inches for household dwellings, commercial, and industrial establishments, respectively. The total freshwater consumption has increased from  $3 \times 10^6 \text{ m}^3$  in the year 1957 to  $160 \times 10^9 \text{ m}^3$  in the year 2018, which represents in 60 years a 53,000 fold increase. Tremendous changes had occurred to the freshwater distribution network over the years, from just 112 kilometers in length in 1959 to over 9,800 kilometers in 2018. Similarly, the brackish water distribution network had increased from 90 kilometers in length in 1959 to over 8,100 kilometers in 2018. Moreover, the number of consumers connected to the freshwater network totaled 177,118 and consumers connected to the brackish water network totaled 77,257 by the end of 2018. However, there were 1,416 and 183 pipe breaks for the freshwater and brackish water, respectively. Figure 1 shows the development of both the fresh and brackish water distribution networks in Kuwait from 1959 to 2018 [18, 19].

Figure 2 displays the total number of pipe breaks for both fresh and brackish water networks during 1999 to 2018. It is noticed that the number of freshwater pipe breaks are by far higher than the brackish water pipe breaks which is due to the higher use and consumption of freshwater daily. In addition, brackish water is available for customers once a week only.

## 2. Model Formulation and Discussion of Results

The Ministry of Electricity and Water (MEW), as well as other water authorities all over the world, is facing a difficult challenge in assessing the physical condition of its distribution systems. Since the majority of the mains are buried, the MEW must rely upon indirect methods, including analysis of repair records. The time to failure of a component is represented as a random variable,  $T > 0$ , with a probability density function  $f(t)$  and a distribution function  $F(t)$ . The

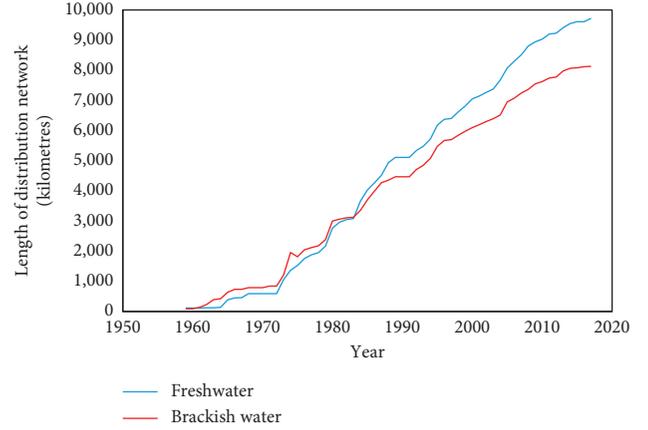


FIGURE 1: Development of fresh and brackish water distribution networks (1959–2018).

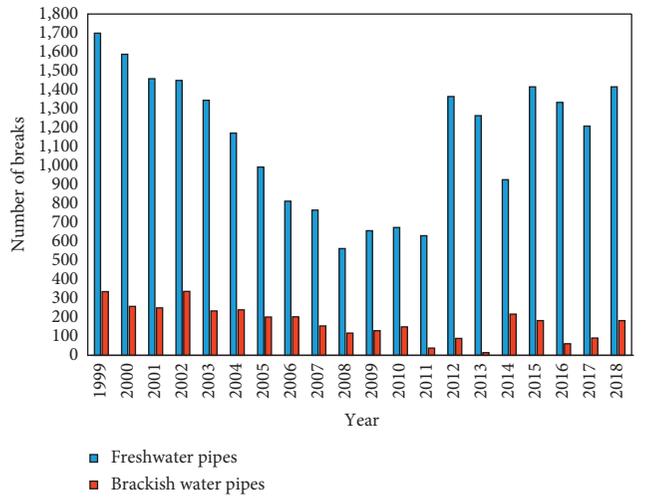


FIGURE 2: Total number of fresh and brackish water pipe breaks.

survival function  $S(t)$  and the hazard function or failure rate  $h(t)$  are defined as [20–22]

$$S(t) = P[T > t] = 1 - F(t), \quad (1)$$

$$h(t)dt = P[t < T < t + dt] = \frac{f(t)dt}{[1 - F(t)]}$$

On integration, one obtains

$$\int_0^t h(u)du = \int_0^t \left\{ \frac{f(u)du}{[1 - F(u)]} \right\} = -\ln[1 - F(t)], \quad (2)$$

$$\int_0^t h(u)du = -\ln S(t).$$

So that

$$S(t) = e^{\int_0^t h(u)du}. \quad (3)$$

The failure time distribution of pipes in a water distribution system may therefore be investigated either through

its survival function  $S(t)$  or its failure rate  $h(t)$ . It is instructive to work with the  $h(t)$  because its behavior can easily be examined in terms of commonly known probability density functions. For example, if  $T$  has an exponential density function, then

$$f(t) = \lambda e^{-\lambda t}, \quad (4)$$

and it is easily seen that

$$S(t) = 1 - F(t) = e^{-\lambda t}. \quad (5)$$

Thus, the failure rate  $h(u) = \lambda = a$  constant. Since, intuitively, we know that the failure rate tends to vary with time, we conclude that the exponential model is not a suitable failure density function.

We would also expect the failure rate to depend on the age of the component as well as other structural and environmental factors. For water mains, experience tells us that some of these factors include age, pipe size and material, internal pressure, and environmental conditions such as cover conditions and soil corrosivity. Studies have been made to relate the failure rate to these explanatory variables. For example, Shamir and Howard [23] used regression analysis to express the failure time as a function of other variables. Their results have been quite unsatisfactory with coefficients of determination (i.e.,  $R^2$ ) usually less than 0.5.

One shortcoming of this type of analysis is that by its nature, the data used in regressing failure time with the other explanatory variables cannot include information on those components that have not yet failed. This suggests that survival analysis which includes censored data may be a more productive technique for analysis. The proportional hazard model is examined in this paper. This model has a hazard function of the form

$$h(t, z) = h_0(t)e^{-zB}, \quad (6)$$

where  $z$  is a row vector of explanatory variables,  $B$  is a column vector of regression parameters, and  $h_0$  is an unspecified baseline hazard function.

This model is attractive for the present problem because the effect of aging is separated from the effects of site-specific causes of failure. This model is proportional because of the multiplicative relation between  $h_0$  and the log-linear function of the explanatory variables. Because of this proportionality assumption, the ratio of the failure rates of two components with different sets of covariates does not depend on time.

The proportional hazard model has been successfully used as a tool for survival analysis in the biomedical field and industrial life testing. Software for survival analysis using the model is available as part of the BMDP statistical package [24]. With an accessible database, the software permits the identification of the significant variables that should be included in the covariate vector. The package also permits the stratification of data in cases where the subsets of the database may follow different patterns.

A desirable feature of the analysis is that it is able to include in the estimation procedure, information of both failed and unfailed components. It thus makes full use of the

information. Survival analysis is in fact distinguished from other fields of statistics through this ability to analyze censored data.

The data used in this study were supplied by the Ministry of Water and Electricity (MWE) and consisted of information on 1,275 pipes in the water distribution system of the Kuwait city. For each pipe, the information included year of installation, length of pipe, diameter, soil corrosivity, material, number and age at break events, internal pressure, cover index, and soil stability. Only 21 percent of the pipes had experienced one or more breaks. Of these, 14.5% had 2 breaks or more and 1.7 percent had been broken 3 or more times. Data are thus censored since 79 percent of the pipes have not failed.

Prior to determining the proportional hazard model for the data set, several analyses were made to determine which explanatory variables are sufficiently significant to be included in the model and to determine if subsets of the data exhibit different failure patterns from other subsets. Procedures for these are described in the BMDP User's Manual. The technique is also discussed by the authors in [6, 25, 26], who had performed similar preprocessing procedures earlier. The results of the prior analysis identified the following variables as significant:

LNLEN: the natural logarithm of the pipe length;

LOWPCT (in decimals): percentage of the pipe in low land use;

PRESS: pipe pressure = kPa if it has not been broken before, or 0 otherwise;

PREBRK: previous break = 2/age at the second break if the previous break equals 2, or 0 otherwise;

BEF70: a dummy variable = 1 if the pipe was installed during the years 1950–70, or 0 otherwise; and

AFT70: dummy variable = 1 if the pipe was installed after 1970, or 0 otherwise.

PRESS was defined as above because it was found that the effect of pressure was not significant after the first break. Similarly, BRKRAT was included as defined because prior analysis suggested that if a pipe breaks often early, it was more prone to break again. BEF70 and AFT70 were included because stratification analysis indicated peculiar break patterns. This may be due to construction practices during these periods.

After the significant variables were identified, fitting the model entailed preparing the input file in the proper format to run the software. A summary of the results includes the value of the coefficient vector  $B$  and the baseline function:

$$h_0(t) = a + bt + ct^2. \quad (7)$$

These are tabulated in the last columns of Tables 1 and 2. The baseline function exhibits the well-known "bathtub" shape typical of many failure patterns.

Another objective of the study was to investigate the effect of censoring on the hazard function  $h(t)$ . In order to do this, several data sets were generated based on sampling the

TABLE 1: Beta values for different cases.

Z	N=1100			N=800			N=500			N=1275
	% censored			% censored			% censored			% censored
	70%	80%	85%	65%	80%	90%	65%	80%	90%	73%
LNLEN	0.39	0.38	0.41	0.46	0.41	0.35	0.46	0.61	0.31	0.42
LOWPCT	-0.60	-0.97	-0.84	-0.67	-0.87	-1.17	0.95	-1.10	-1.64	-0.65
BEF70	-0.55	-0.39	-0.18	-0.44	-0.41	-0.42	-0.41	-0.37	-0.89	-0.51
AFT70	0.59	0.85	1.05	0.65	0.88	1.68	0.86	1.10	1.78	0.59
PREBRK	1.39	1.97	2.68	1.14	2.08	3.10	1.01	2.01	2.50	1.63
PRESS	-0.02	-0.02	-0.01	-0.01	-0.01	-0.03	-0.02	-0.01	-0.07	-0.01
BRKRAT	0.71	0.69	0.76	0.68	0.70	0.72	0.67	0.75	0.71	0.70

TABLE 2: Fitted coefficients for baseline hazard functions\*.

	N=1100		N=800		N=500		N=1275
	% censored		% censored		% censored		% censored
	70%	85%	65%	90%	65%	90%	73%
<i>a</i>	0.0191	0.0045	0.02458	0.0014	0.02442	0.0012	0.01405
<i>b</i>	-0.0012	-0.0003	-0.00148	-0.0002	-0.00148	0.00001	-0.0009
<i>c</i>	0.00002	0.0000	0.000019	0.0000	0.00029	0.00000	-0.00001

\*( $h_0 = a + bt + ct^2$ ).

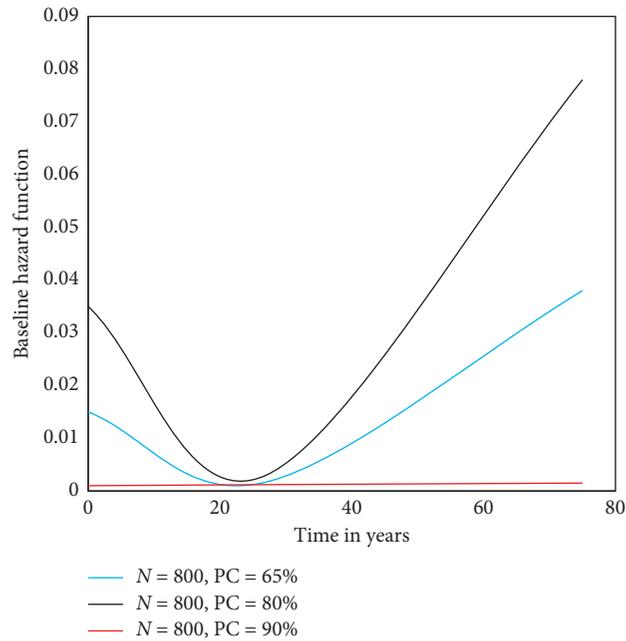


FIGURE 3: Baseline hazard function (PC-variable).

given database. Through a random sampling procedure, additional databases were developed with values of  $N=600$ , 900, and 1200, each having a different percentage of censored observations. These 9 simulated data sets were then analyzed as before to develop the hazard function and to examine how sensitive are the regression coefficients to sample size and percentage of censored observations. The results are important in deciding when to terminate a

sampling program or how big a sample size must one have. The results are shown in Tables 1 and 2. Figure 3 displays the  $h_0(t)$  for  $N=800$  and different percentages of censored data. Figure 4 shows  $h_0(t)$  for different  $N$ s with the same censored data percentage. Moreover, the other advantage of this study is that the most important parameters (regarding pipe failures and breaks) for maintenance and repair decisions are identified.

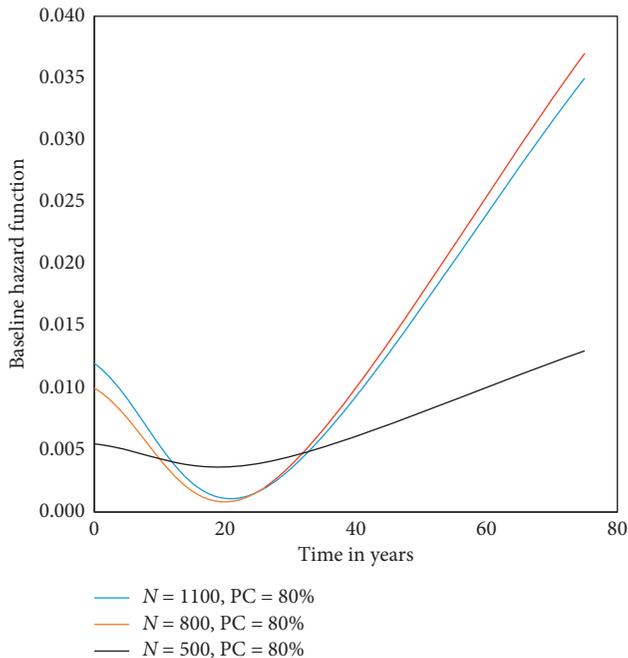


FIGURE 4: Baseline hazard function ( $N$ -variable).

### 3. Conclusions

The Ministry of Water and Electricity (MEW) in Kuwait has accumulated huge amounts of water pipe data over the years. However, most of the historical data kept in files have very limited applications. This study develops in detail the procedures for converting voluminous data in files into very useful and reliable information. Pipe failures and breaks in Kuwait were analyzed and modeled through the proportional hazard model. Parameters (regarding pipe failures and breaks) that are useful for maintenance and repair decisions are identified. It is shown that the proportional hazard model is suitable for describing failure rates of components. The hazard function varies with sample size and the percentage of censored data. This variability must be kept in mind when making inferences. It has the desirable feature of being able to separate the time-dependent effect of aging from other stress factors which are often site specific. Another advantage is its suitability for analyzing censored data, and thus can efficiently use information from a database. The availability of the software and the facility of its use is another advantage. Moreover, its ability to stratify nonhomogeneous data and identify the important explanatory variables make it suitable for failure modeling not only in the water distribution network but also in other constructed systems, such as sewer and gaslines, where it can give valuable insights about such systems.

### Data Availability

All the data used are available in the manuscript and can also be requested from the authors.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

The authors are thankful to the Ministry of Electricity and Water (MEW) for supplying the relevant data.

### References

- [1] V. Kanakoudis and S. Tsitsifli, "Water pipe network reliability assessment using the DAC method," *Desalination & Water Treatment*, vol. 33, no. 1–3, pp. 97–106, 2011.
- [2] K. Pietrucha-Urbanik and A. Zelazko, "Approaches to assess water distribution failure," *Period Polytech-Civ*, vol. 61, pp. 632–639, 2017.
- [3] J. R. Rak and K. Pietrucha-Urbanik, "An approach to determine risk indices for drinking water-study investigation," *Sustainability*, vol. 11, no. 11, p. 3189, 2019.
- [4] S. Tsitsifli and V. Kanakoudis, "Predicting the behavior of a pipe network using the "critical Z-score" as its performance indicator," *Desalination*, vol. 250, no. 1, pp. 258–265, 2010.
- [5] S. Tsitsifli, V. Kanakoudis, and I. Bakouros, "Pipe networks risk assessment based on survival analysis," *Water Resources Management*, vol. 25, no. 14, pp. 3729–3746, 2011.
- [6] G. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden Day, San Francisco, CA, USA, 1976.
- [7] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 34, no. 2, pp. 187–202, 1972.
- [8] W. A. Fuller, *Introduction to Time Series*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2nd edition, 1996.
- [9] V. K. Kanakoudis and D. K. Tolikas, "Assessing the performance level of a water system," *Water, Air & Soil Pollution*, vol. 4, pp. 307–318, 2004.
- [10] V. K. Kanakoudis, "Vulnerability based management of water resources systems," *Journal of Hydroinformatics*, vol. 6, no. 2, pp. 133–156, IWAP, 2004.
- [11] V. K. Kanakoudis and D. K. Tolikas, "The role of leaks and breaks in water networks: technical and economical solutions," *Journal of Water Supply: Research and Technology-Aqua*, IWAP, vol. 50, no. 5, pp. 301–311, 2001.
- [12] S. A. Andreou, D. H. Marks, and R. M. Clark, "A new methodology for modelling break failure patterns in deteriorating water distribution systems: Theory," *Advances in Water Resources*, vol. 10, no. 1, pp. 2–10, 1987.
- [13] U. Shamsi, "GIS application for water distribution systems," *Journal of Water Management Modeling*, pp. 459–472, 2004.
- [14] R. A. Francis, S. D. Guikema, and L. Henneman, "Bayesian Belief Networks for predicting drinking water distribution system pipe breaks," *Reliability Engineering & System Safety*, vol. 130, pp. 1–11, 2014.
- [15] U. Shamsi and R. G. Quimpo, *Reliability Approach to Water Supply Infrastructure Maintenance*, Research Report CE/WR-01, University of Pittsburgh, Pittsburgh, PA, USA, 1988.
- [16] S. Yamijala, S. D. Guikema, and K. Brumbelow, "Statistical models for the analysis of water distribution system pipe break data," *Reliability Engineering & System Safety*, vol. 94, no. 2, pp. 282–293, 2009.
- [17] J. M. Alhumoud, "Freshwater consumption in Kuwait: analysis and forecasting," IWA, *Journal of Water Supply: Research and Technology-AQUA*, vol. 57, 2008.
- [18] MEW, *Ministry of Electricity and Water, Statistical Department and Information Center (Water)*, State of Kuwait, MEW, Kuwait, State of Kuwait, 2019.
- [19] F. Al-Ruwaih, J. Alhumoud, and S. Al-Mutairi, "Quality of potable water in Kuwait," *American Journal of Environmental Sciences*, vol. 6, no. 3, pp. 260–267, 2010.

- [20] A. Birolini, *Reliability Engineering, Theory and Practice*, Springer-Verlag GmbH Deutschland, Berlin, Germany, 8th edition, 2017.
- [21] R. Billinton and R. N. Allan, *Reliability Evaluation of Engineering Systems*, Plenum Press, New York, NY, USA, 1983.
- [22] M. E. Harr, *Reliability Based Design in Civil Engineering*, McGraw Hill Book Company, New York, NY, USA, 1987.
- [23] U. Shamir and C. P. D. Howard, "An analytical approach to pipe replacement," *Journal of the American Water Works Association*, vol. 71, no. 5, 1979.
- [24] BMDP, *Statistical Software Manual*, University of California Press, Los Angeles, CA, USA, 1995.
- [25] J. E. Shortridge and S. D. Guikema, "Public health and pipe breaks in water distribution systems: analysis with internet search volume as a proxy," *Water Research*, vol. 53, pp. 26–34, 2014.
- [26] Q. Xu, Q. Chen, W. Li, and J. Ma, "Pipe break prediction based on evolutionary data-driven methods with brief recorded data," *Reliability Engineering & System Safety*, vol. 96, no. 8, pp. 942–948, 2011.