Forecasting Failure of Water Mains Under Climate Variations: Stochastic Modeling Process

Zainab Almheiri[1,4], Mohamed Meguid[2] & Tarek Zayed[3]
[1] PhD Candidate, Dept. of Civil Engineering and Applied Mechanics, McGill University, Montreal, Quebec, Canada.
[2] Professor and Chair, Dept. of Civil Engineering and Applied Mechanics, McGill University, Montreal, Quebec, Canada.
[3] Professor, Building and Real Estate Dept (BRE), The Hong Kong Polytechnic University, Hong Kong, China.
[4]E-mail address: zainab.almheiri@mail.mcgill.ca

ABSTRACT
Climate change has the potential to affect substantially the operational condition and lifespan of buried utilities, including pipelines transporting water, gas, and sewer. The climate of Canada is known to vary spatially and temporally, and these variations can cause damage to buried infrastructures. The physical mechanisms that negatively impact the condition of water mains and can lead to their failure are complex and not fully understood. High uncertainty could result from forecasting the potential damage and the need to replace or repair water mains before catastrophic failures. Although climate risks have become a significant concern to cities and municipalities in Canada, only a few studies have addressed the impact of climate variations on the failures of water mains. In this study, we present a stochastic modeling process for failure forecasting of water mains at selected locations in the Cities of London (Ontario) and Quebec, Canada.

RÉSUMÉ
Le changement climatique a le potentiel d'affecter considérablement la condition opérationnelle et la durée de vie des services publics enterrés, y compris les pipelines transporteurs d'eau, de gaz et d'égouts. Le climat du Canada est connu par sa variation dans l'espace et dans le temps. Ces variations peuvent endommager les infrastructures enfouies. Les mécanismes physiques qui impactent négativement l'état des conduites d'eau et qui peuvent entraîner leur défaillance sont complexes et ne sont pas entièrement compris. Une incertitude élevée pourrait résulter de la prévision des dommages potentiels et de la nécessité de remplacer ou de réparer les conduites d'eau avant les pannes catastrophiques. Bien que les risques climatiques soient devenus une préoccupation importante pour les villes et les municipalités du Canada, très peu d'études ont adressé l'impact des variations climatiques sur les défaillances des conduites d'eau. Notre étude vise à présenter un processus de modélisation stochastique pour la prévision des défaillances des conduites d'eau à des endroits sélectionnés dans les villes de London (Ontario) et Québec au Canada.

1 INTRODUCTION

Water mains are critical infrastructures, and their failure can cause long-term catastrophic damage to urban water distribution infrastructures. Additionally, wasting a large amount of treated water (Almheiri et al., 2020). Deterioration process of water mains is complex. Many factors may contribute to the failure of water mains including static, dynamic, and operational (Haider, Sadiq, & Tesfamariam, 2013; Rajani & Tesfamariam, 2004).

The contribution of static and operational factors on the failure of water mains have studied widely by previous researchers. Jafar et al. (2010) utilized artificial neural networks (ANN) to predict the failure rate of water mains based on operational and static variables. Francis et al. (2014) developed Bayesian belief networks (BBNs) to predict the failure of water mains using pipe materials, age, diameter, demographic variables, and temperature. Additionally, Shirzad et al. (2014) applied the support vector machine (SVM) and ANN to predict the failure rate of water mains. Almheiri et al. (2020) developed three intelligent approaches to predict time-to-pipe failure based on static factors.

Climate changes and temperature variations may contribute to the failure of water mains (Boyle, Cunningham, & Dekens, 2013). Vreeburg et al. (2013) reported that the number of failures of water mains increases during the winter season. Similarly, Bruaset and Saegrov (2018) reported that the failure rate of water mains increases with the decrease in temperature.

Frost loading, on the other hand, can impose significant strains on pipe walls which lead to increase the chance of failure (Bruaset & Saegrov, 2018). Soil movement due to thermal expansion and shrinkage during seasonal changes can result in pipeline failure (Bruaset & Saegrov, 2018).

Previous studies have addressed the impact of temperature variations on the failure of water mains. However, limited studies have examined the impact of climate variations on the failure prediction of water mains. Forecasting failure of water mains is essential in maintaining the sustainability of water distribution systems.

To fill this gap, we propose a stochastic modeling process to forecast the failure of water mains under climate variations at selected spatial locations. In this study, we focus on climate variables such as temperature, freezing index, frost days, and dry index. To achieve this, we collect spatiotemporal data to study the failure behavior of water mains at selected spatial locations over a defined period from 1987 to 2001. The proposed framework can capture periodic and temporal patterns of water main at selected spatial locations.

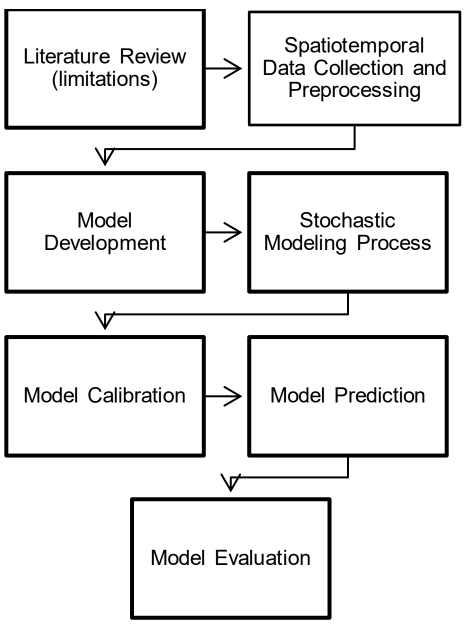


Figure 1. Proposed framework of forecasting failure of water mains under environmental and climate variations.

2 METHOD

2.1 Spatiotemporal Data

Spatiotemporal data are collected from different spatial locations in Canada, Quebec and London Ontario. Spatiotemporal data manage both time and space information. Climate data are collected from the Environment and Climate Change Canada (ECCC) at selected spatial locations. Failure data of water mains, on the other hand, are obtained from Sainte-Foy and the City of London municipalities.

Failure data consist of failure frequency of water mains over a defined period at selected spatial locations, from 1987 to 2001. Whereas, climate data consist of temperature in Celsius (ºC), and precipitations (mm). Additionally, freezing index in Degree-days (°Cd) and frost days (Days) are calculated to study their impact on the failure of water mains.

Freezing index is a measure of winter severity and defined as the total average daily temperature below zero (see Eq.1). Similarly, frost days are defined as the total days where the daily average temperature is below zero. Dry index, on the other hand, is used as a surrogate measure of soil moisture (Eq. 2), (Fuchs-Hanusch et al., 2013).

$$Freeding\ Index_{(i)} = \left|\sum_{i=1}^{n} \bar{T}\right| , \bar{T} < 0. \quad [1]$$

$$Dry\ Index = \frac{3 \times \bar{T}}{P}. \quad [2]$$

2.2 Experimental Setup

Data preprocessing is a paramount step in most of the learning algorithms. First, we associate the given failure frequency of water mains per month per year with climate data through the date of the failure of water mains over the studied period (from 1987 to 2001). Then, we apply data normalization to all variables to be in the range of [0, 1]. Finally, we present the variables in a matrix form with size of $t \times d$, where t is the length of sequential climate variables, and d is the number of climate variables.

$$\begin{bmatrix} \boldsymbol{x_{11}} & \cdots & \boldsymbol{x_{1d}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{x_{t1}} & \cdots & \boldsymbol{x_{td}} \end{bmatrix}$$

2.3 Stochastic Modeling Process

Time series at different spatial locations can be considered as multivariate time series. The interdependency between multivariate time series can be studied by developing time series models. Vector autoregression (VAR) is a stochastic process model that links the values of the multivariate time series at a time *t* to some linear combinations of the multiple time series at previous times. Suppose a K-variate time series VAR(p) model is given by:

$$\boldsymbol{Y_t} = \boldsymbol{\Phi_1 Y_{t-1}} + \boldsymbol{\Phi_2 Y_{t-2}} + \cdots + \boldsymbol{\Phi_p Y_{t-p}} + \boldsymbol{W_t}\,, \quad [3]$$

where $\boldsymbol{Y_t} \equiv (Y_t^{(1)}, \ldots, Y_t^{(K)}), \{\boldsymbol{\Phi}_l : l = 1, \ldots, p\}$ are $K \times K$ state-transition matrices and $\{\boldsymbol{W_t}\}$ is a white-noise process for each K-variate with a mean of zero and a covariance matrix **Q.**

In this study, we propose a VAR modeling process with exogenous climate variables to forecast the failure of water mains under climate variations at selected spatial locations. Eq. 4 predicts the failure of water mains $\mathbf{Y_t}$ at selected spatial locations based on its previous p values and climate variables $\mathbf{X_t}$.

$$Y_t = \sum_{l=1}^{p} \boldsymbol{\Phi}_l Y_{t-l} + \sum_{j=1}^{s} \boldsymbol{B}_j \boldsymbol{X}_{t-j} + \boldsymbol{W}_t, \quad [4]$$

Where s values of a stationary m-dimensional vector series $\mathbf{X_t}$, and B $\{\mathbf{B}_j \in R_j^{k \times k}\}_{j=1}^{s}$ are the estimated exogenous parameter matrices.

2.4 Model Calibration

Model calibration is a paramount step in all mathematical modeling to assure numerical stability and develop an accurate model. Many parameters are selected at the beginning and reduced to avoid model complexity that leads to overfitting.

Time lags are the hyperparameter for the proposed model. The best model among several candidate models was then selected based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC) to develop the final model (see Eq. 5 and Eq. 6).

$$\boldsymbol{AIC} = -\boldsymbol{2}\,\boldsymbol{log}\,\boldsymbol{L}(\hat{\boldsymbol{\theta}}) + \boldsymbol{2k}. \quad [5]$$

$$\boldsymbol{BIC} = -\boldsymbol{2}\,\boldsymbol{log}\,\boldsymbol{L}(\hat{\boldsymbol{\theta}}) + \boldsymbol{2}\,\boldsymbol{log}\,\boldsymbol{k}. \quad [6]$$

where θ the model parameter, and k the number of estimated parameters of the candidate model.

2.5 Model Evaluation

The developed model is evaluated using a set of mathematical equations. Prediction mean square error (PMSE), root mean square error (RMSE), and accuracy (Eqs. 7 to 9).

$$\boldsymbol{PMSE} = \frac{\boldsymbol{1}}{\boldsymbol{T}} \sum_{i=1}^{T} (\boldsymbol{FF}(\boldsymbol{i}) - \boldsymbol{forecast.FF}\,(\boldsymbol{i}))^2. \quad [7]$$

$$\boldsymbol{RMSE} = \sqrt{\frac{\boldsymbol{1}}{\boldsymbol{T}} \sum_{i=1}^{T} (\boldsymbol{FF}(\boldsymbol{i}) - \boldsymbol{forecast.FF}\,(\boldsymbol{i}))}. \quad [8]$$

$$\boldsymbol{Accuracy} = \boldsymbol{1} - \frac{\boldsymbol{1}}{\boldsymbol{T}} \sum_{i=1}^{T} \left| \frac{\boldsymbol{forecast.FF}\,(\boldsymbol{i}) - \boldsymbol{FF}(\boldsymbol{i})}{\boldsymbol{FF}(\boldsymbol{i})} \right|. \quad [9]$$

*where FF refers to the failure frequency at selected spatial locations.

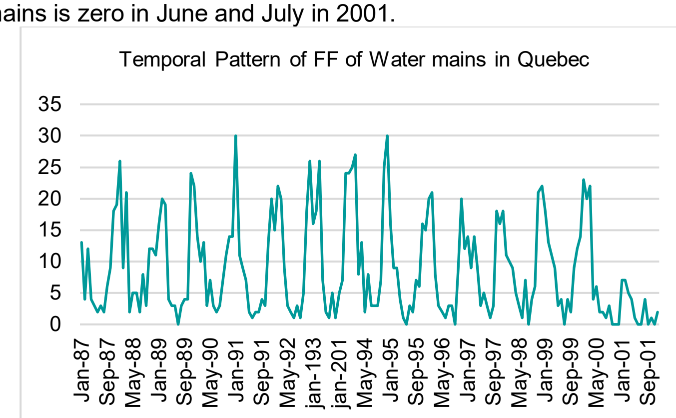
3 RESULTS

3.1 Failure Frequency

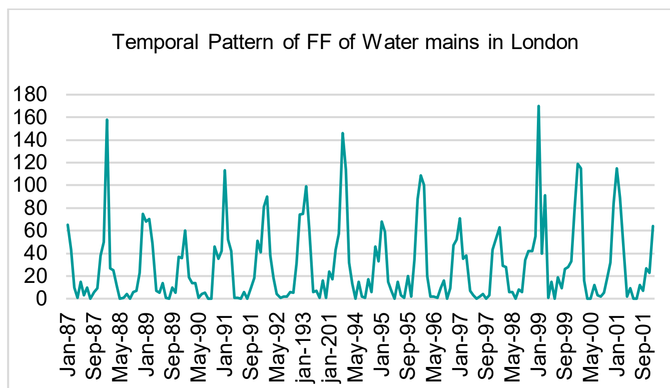
The failure of water at selected spatial locations (Quebec and London Ontario) follows a periodic pattern, as shown in Figure 2. The failure frequency of water mains in London is higher than in Quebec.

The number of failure frequency of water mains at selected locations occur in all seasons, winter, summer, spring, and fall. However, the chance of failure increases during winter seasons.

Table 1 illustrates the monthly failure frequency at selected spatial locations in 2001. Similarly, to the whole defined period, the failure frequency of water mains reached its highest value in January at both selected spatial locations. Whereas, the failure frequency of water mains is zero in June and July in 2001.



(a)



(b)

Figure 2. Temporal pattern of failure frequency of water mains at selected spatial locations.

Table 1. Monthly failure frequency (FF) at selected spatial locations in two cities in Canada in 2001.

*highest, **lowest FF occurred during wintertime (January) and summertime (June-July) over the defined period.

Total FF	Quebec	London
January*	7	115
February	7	89
March	5	41
April	4	2
May	1	9
June**	0	0
July**	0	0
August	4	12
September	0	7
October	1	27
November	0	23
December	2	64

3.2 Model Performance

The model is developed and then tested on held-out data to verify the model performance. The results prove that the model can capture the periodic and temporal pattern of the failure of water mains. The model can also predict the failure of water mains up to nine months ahead under climate variations. The average prediction accuracy of the developed model is 81%.

The failure process of water mains is complex, and it relied on many other factors besides climate variations. Thus, the proposed model is designed to accommodate additional variables that may contribute to the failure of water mains at different spatial locations.

4 CONCLUSION

In this study, we propose a stochastic modeling process that can predict the failure of water mains at selected spatial locations under climate variations. The model, on the other hand, can be extended to accommodate other variables that contribute to the failure of water mains.

Through data exploration, we find that the failure of water mains at selected spatial locations follows a periodic temporal pattern. Besides, the failure of water mains may occur during any season, however, it drastically increases during winter.

5 ACKNOWLEDGEMENTS

This research is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Financial support provided by McGill-UAE fellowships in Science and Engineering to the first author is highly appreciated.

6 REFERENCES

Almheiri, Z., Meguid, M., & Zayed, T. (2020). Intelligent Approaches for Predicting Failure of Watermains. *ASCE Journal of Pipeline Systems Engineering and Practice* (In press).

Boyle, J., Cunningham, M., and Dekens, J. (2013). Climate Change Adaptation and Canadian Infrastructure. *International Institute for Sustainable Development (IISD)*: Winnipeg, MB, USA.

Bruaset, S., and Saegrov, S. (2018). An Analysis of the Potential Impact of Climate Change on the Structural Reliability of Drinking Water Pipes in Cold Climate Regions. *Water*, 10(4), p. 411.

Fuchs-Hanusch, D., Friedl, F., Scheucher, R., Kogseder, B., and Muschalla, D. (2013). Effect of seasonal climatic variance on water main failure frequencies in moderate climate regions. *Water Science and Technology*: Water Supply, 13(2), pp. 435–446.

Francis, R. A., Guikema, S. D., and Henneman, L. (2014). Bayesian belief networks for predicting drinking water distribution system pipe breaks. *Reliability Engineering & System Safety*, 130, pp. 1–11.

Haider, H., Sadiq, R., and Tesfamariam, S. (2013). Performance indicators for small-and medium-sized water supply systems: a review. *Environmental Reviews*, 22(1), pp. 1–40.

Jafar, R., Shahrour, I., & Juran, I. (2010). Application of Artificial Neural Networks (ANN) to model the failure of urban water mains. *Mathematical and Computer Modelling,* 51(9-10), 1170-1180.

Rajani, B., and Tesfamariam, S. (2004). Uncoupled axial, flexural, and circumferential pipe soil interaction analyses of partially supported jointed water mains. *Canadian Geotechnical Journa*l, 41(6), pp. 997–1010.

Shirzad, A., Tabesh, M., & Farmani, R. (2014). A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks. *KSCE Journal of Civil Engineering*, 18(4), 941-948.

Vreeburg, J., Vloerbergh, I., Van Thienen, P., and De Bont, R. (2013). Shared failure data for strategic asset management. *Water Science and Technology: Water Supply*, 13(4), pp. 1154–1160.