"6th International Conference on Water Economics, Statistics and Finance and 10th Leading Edge Conference for Strategic Asset Management (LESAM)"
Pafos, Cyprus, 28-30 April, 2025

Sewerage Report Analysis for Report-Based Failure Prediction in Pipes

Ofek Aloni[1], Eran Friedler[2], Daniel Sauter[4], Lukas Guericke[3], Nicolas Caradot[3], Barak Fishbain[1,2*]

[1]*Department of Mathematics, Technion, Haifa, Israel*
[2]*Department of Civil and Environmental Engineering, Technion, Haifa, Israel*
[3]*KWB Kompetenzzentrum Wasser Berlin gemeinnützige GmbH, Berlin, Germany*
[4]*Berliner Wasserbetriebe, Neue Jüdenstraße 1, 10179 Berlin, Germany*
**Corresponding author: fishbain@cv.technion.ac.il*

Maintenance of sewerage networks is costly, and the funds required will only increase with population size and pipe aging. Smart tools for future planning can help reduce costs and focus repairs on the pipes that need them most, preventing failures. Early research shows promise of Machine Learning (ML) models such as Random Forests (RF) in predicting the evolution in time of the structural state of pipes (Caradot et al., 2018). We take a different approach, using ML models to directly predict pipe failures. Training data utilizes past reports of pipe failures, combined with GIS data with pipe parameters. This makes it immediately applicable to many municipalities, as this data is often available. This is especially relevant when CCTV inspections assessing the structural condition of pipes are scarce.

Data was obtained from two water and sewer corporations in the Sharon area of the country, namely Mey Natanya and Maayanot Hasharon. GIS data was obtained from each, containing information regarding all pipes in the system. Included in the data were pipe location and length, as well as pipe properties, such as age, material, and diameter, among others. The other set of data supplied is reports, tabular data of citizen complaints spanning about 10 years. The tabular reports were filtered to exclude reports that do not necessarily indicate failures, e.g. complaints of smell, or misplaced manhole covers. Remaining reports were geolocated and matched to the closest pipe.

Exploratory data analysis was conducted on the processed reports. Initially, heatmaps and animated occurrence of reports were generated, to look for areas of interest. The monthly distribution of reports was plotted to observe seasonal trends. Indeed, a higher occurrence of reports was found during the wet season.

A system-level analysis that was of interest to us was recurring reports in pipes. As seen in Figure 1, both systems seem to obey exponential decay when applying ln and normalizing. That led us to a conjecture that γ, the coefficient of decay, might characterize the system, as it is significantly different in the different municipalities. Given more samples of similar data, γ might show correlation with the properties of the pipes in the system, e.g. typical age of pipes, average pipe length, etc. Alternatively, it could be tied to system-wide properties, i.e. total number of pipes.

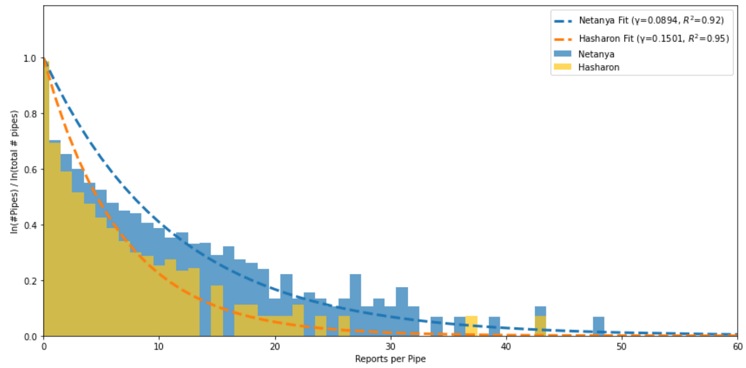

Figure 1 – Log$_e$ of the number of pipes with a certain number of reports, with exp. fit for γ

Formatting the report data as yearly binary data, a temporal analysis of the reports was conducted. The conditional probability for failure was calculated for different conditions on the years before. Namely, the persistence – having a report in each of the previous n years, and the conditional probability of a report, given a report in the previous year, as well as reports in at least one of the other last n years:

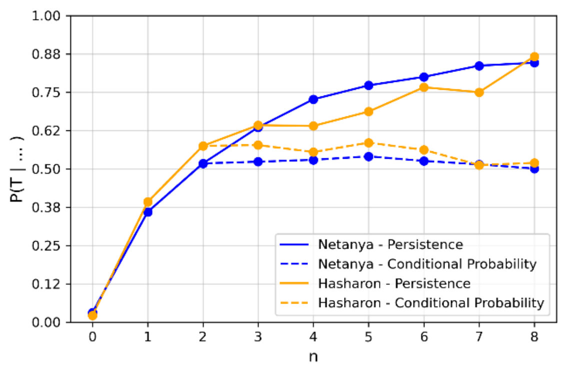

Figure 2 - Probabilities for report - persistence (solid) and conditional probability (dashed), both as function of the number of lags n

This abstract highlights efforts to predict sewerage system failures using complaints report data, covering data processing, analysis insights, and initial modelling. Next steps include testing graph neural network (GNN) architectures to boost performance and exploring links between report data and infrastructure conditions, requiring additional network data.

Acknowledgement: This work is partially funded by the German Federal Ministry of Education and Research (BMBF) and Israel Ministry of Science and Technology (MOST) as part of the project “DASAM: Data-driven Sewer Asset Management in Germany and Israel” (funding code 02WIL1700A). The responsibility for the content of this publication lies with the authors.

Key words: Predictive maintenance, Sewage, Wastewater systems

References

Caradot, N., Riechel, M., Fesneau, M., Hernandez, N., Torres, A., Sonnenberg, H., Eckert, E., Lengemann, N., Waschnewski, J., & Rouault, P. (2018). Practical benchmarking of statistical and machine learning models for predicting the condition of sewer pipes in Berlin, Germany. *Journal of Hydroinformatics*, *20*(5), 1131–1147. https://doi.org/10.2166/hydro.2018.217