infraPLAN
Water Mains R&R & AI

Ask the Experts! #3
Superior LoF Score

Click below to read all ATE articles.
Article 1: 1pc Water Main R&R Rate
Article 2: LoF Scores
Article 3: LoF Case Studies
Article 4: Data Needed
Article 5: Abandoned Pipes
Article 6: Missing Data
Article 7: Structural Data Issues

Break Prediction using infraSOFT Machine Learning Module is up to 6 timesSuperior to Desktop Scoring

Case studies

In a previous article (#2) we have shown that assessing the Likelihood of Failure (LoF) of a pipe using desktop or GIS scoring can be difficult to complete and limited in scope; and that relying on advanced analytics is easier to achieve.

In this article we compare the break prediction capacity of desktop scoring, multi-variable regression, powered or not by machine learning when applied to three systems of different size, physical condition, data quality, and break history.
We show that the capacity of our proprietary machine learning model -embedded in our **infraSOFT** platform- to predict breaks is up to 6 times superior to desktop scoring; up to 90% of the breaks occurring during the next two years can be predicted. Our machine learning model performs extremely well even on a small system, or a system with an erratic break trend, or problematic data.

Our machine learning-powered break prediction model

Machine learning essentially pushes data, and allows various human tasks to be conducted automatically based on a pre-defined logic. It makes the calibration of statistical models more systematic. However, the model used to analyze the data (and establish the multiple connections between all the variables) still needs to be defined by the analyst. With this respect, we came to the field of predicting breaks using machine learning with 15 years of experience developing multi-variable advanced statistical regression models (Weibull, Cox, Proportional Hazard Model, and more recently, Linear Extended Yule Process).

We understand the data used, and the constraints imposed to the models fed to the machine. For example, we ensure the model be of a survival nature, i.e., that abandoned pipes be taken into account. In the three case studies that follow, we compare the predictive capacity of up to three models: desktop scoring, multi-variableregression (LEYP) carefully calibrated manually, and our own statistical model powered this time by machine learning. All the predictable steps of a calibration are handled automatically when machine learning is used. There may always be some need for unforeseeable manual calibration even with machine learning.

The validation technique is described in the first case study; it is the same regardless of the modeling approach.

Project #1 - Large system - Medium Break Rate
1,490.0 miles of Cast Iron (CI) pipes and 713.6 miles of Ductile Iron (DI) pipes

Project #1 took place in 2022 with break data from 2002-2021. Breaks from 2002-2019were used to generate a predicted LoF with each modeling approach considered -for this project, *desktop scoring* and *machine learning*- for 2020-2021, for each pipe.

The pipes were then ranked based on their LoF, with the highest LoF first. For each pipe we also have the actual breaks that occurred in 2020-2021. We consider a certain percentage of pipes (ranked with highest LoF first) and compute the percentage of the total number of breaks experienced by those pipes in 2020-2021; the idea is that the breaks would have been avoided had those pipes been replaced by 2020. The higher the percentage of "breaks avoided" during the validation period, the better the model.

The table below shows the validation results obtained with *desktop scoring* and *machine learning* for Project #1. *Multi-variable regression* was not used for this project.

ALL PIPES	% Breaks (2020-2021) avoided (worst pipes first)		
% Pipes Targeted	Desktop scoring	Multi-variable regression	Machine Learning
1	17.8	NA	30.6
5	**19.4**	NA	**79.3**
10	**45.2**	NA	**84.9**

The table reads as follows: (second row) if the top **5%** worst pipes, as ranked by their *desktop scoring* LoF score, had been replaced by 2020, **19.4%** of the 2020-2021 breaks would have been avoided; if ranked with *machine learning* LoF scores, that percentage becomes **79.3%.** For the top **10%** worst pipes, the percentages are **45.2%** (desktop) versus **84.9%** (ML). The performance of the *machine learning* model is excellent with this data set.

For project #1, the capacity of our machine learning model to predict breaks is 1.8 to 4 times better than desktop scoring.

Project #2 - Very Large system - High break rate
2,457.7 miles of Cast Iron (CI) pipes and 2,511.7 miles of Ductile Iron (DI) pipes.

Project #2 took place in 2021 with 2005-2020 break data; 2005-2018 breaks were used to predict breaks for 2019-2020 which were compared with the actual breaks.

The table below shows the validation results obtained with *desktop scoring*, *multi-variable regression*, and *machine learning*. For this project a *multi-variable regression* model had to be created for each material because of calibration constraints that characterize that model (one of the limitations). Therefore, to compareapples to apples, the *desktop scoring* and *machine learning* models were also validated separately for each material.

Each cell contains two values, first the percentage of breaks avoided by DI pipes, andthen by CI pipes.

DI/CI	% Breaks (2019-2020) avoided (worst pipes first)		
% Pipes Targeted	Desktop scoring	Multi-variable regression	Machine Learning
1	0.6/1.1	34.0/12.1	44.2/15.3
5	**12.2/5.9**	**54.5/35.7**	**63.5/38.1**
10	20.2/9.2	69.4/52.0	72.0/55.2

Results show that, for example, if the top **5%** worst pipes, as ranked with LoF scores generated with *desktop scoring*, *multi-variable regression*, or *machine learning*, had been replaced,
- **12.2%, 54.5%,** and **63.5%** of the 2019-2020 breaks would have been avoided on DIpipes, and;
- **5.9%, 35.7%,** and **38.1%** of the 2019-2020 breaks would have been avoided on CIpipes.

For this project, *advanced analytics* also yielded much better results than *desktop scoring*. However, the scale of the improvement depends on the percentage of pipesreplaced, the material and approach.

CI pipes constitute a cohort for which predictions have been more difficult to make regardless of the approach (the data needs further improvement). For those pipes, *desktop scoring* yields results that, at 5%, are just slightly better than rolling dice! However, even for that weak cohort, results are still around 6 timesbetter at 5% and 10% with *machine learning* than with *desktop scoring*.

For project #2, the machine learning break prediction performs up to 6 times better than desktop scoring, even for a set of rather poor data.

Similarly to Project #1, Project #2 also illustrates that advanced analytics has a much higher failure forecasting capacity than desktop scoring, with *machine learning* performing slightly better than *multi-variable regression*.

The good performance of the regression model is partly thanks to adequate calibration choices which have to be made manually by the user for every run. This can be tedious and requires expertise. While a strong *machine learning* model also requires that the specificity of pipe and break data be taken into account, this is embedded in the software; the model is internally programmed by the data scientist for automated calibration.

Project #3 - Small system - Erratic break trend
15.5 miles of Cast Iron pipes and 187.7 miles of Ductile Iron pipes

This study took place in 2020 with 2007-2019 break data. 2005-2017 breaks were used to predict breaks for 2018-2019 which were compared with the actual 2018-2019breaks. Breaks have been erratic with an overall upward trend from 2007 to 2019.

However, spikes were observed in 2010, 2012 and 2015 with up to 4 times more breaks than the previous year (due to change in operations). The table below shows the validation results obtained with *machine learning*; the only model developed for this project.

ALL PIPES	% Breaks (2018-209) avoided (worst pipes first)		
% Pipes Targeted	Desktop scoring	Multi-variable regression	Machine Learning
1	NA	NA	50.0
5	NA	NA	73.3
10	NA	NA	90.0

If the top 1% worst pipes, as ranked by their *machine learning* LoF, had been replaced, 50.0% of the 2018-2019 breaks would have been avoided; for the top 5%, and 10% worst pipes, the percentages are 73.3%, and 90%, respectively.

Project #3 illustrates the fact that even for a small system (less than 200 miles) with an erratic break pattern, machine learning yields excellent break predictions; up to 90% of the breaks for the next two years are predicted.

Click below to read all ATE articles.
Article 1: 1pc Water Main R&R Rate
Article 2: LoF Scores
Article 3: LoF Case Studies
Article 4: Data Needed
Article 5: Abandoned Pipes
Article 6: Missing Data
Article 7: Structural Data Issues

Contact us for a free discussion on using advanced analytics to maximixe your R&R plan!

infraPLAN-LLC helps water utilities, large and small, achieve savings on CIP expenses using our ground breaking platform, **infraSOFT**

www.infraPLAN-llc.com
(917) 349-6386

Annie Vanrenterghem, PhD, CEO