infraPLAN Ask the Experts! #4

Title of email. you may be underestimating how soon your pipes will break. Here is why

Not including abandoned pipes in your break forecasting analysis creates a bias over time

Let's first illustrate our point with an example. We want to predict what the condition of a 60-yr old pipe will be 20 years from now, at 80. We look at the behavior of all similar pipes that ever reached 80 during the period of breaks observation, those still active and those abandoned. However, "similar" pipes are not exactly similar; some are better, some are worse. Let's imagine that the abandoned pipes happen to be the worst in their group (a reasonable assumption as the worst pipes with the most breaks tend to be abandoned in priority.) If the records of those abandoned (worst) pipes were not available, we would be predicting the behavior of the 60-yr-olds in 20 years (that, on another end, still include good and bad pipes) based on the behavior of the 80-year-olds that are left, which are the best ones. This creates a bias; those pipes will be predicted to be in better condition; their lives will be overestimated.

Would it make sense to predict human longevity by including health data solely on the living or healthy population? Would we know about the risk factors of an epidemic if we only included in our study those who have survived?

Statistical analyses pertaining to pipes degradation borrow a lot to human epidemiology. Both fields study a population (of pipes in the water industry/individuals in public health)

over a certain period (of their service/life) and look for certain occurrences (break/disease) having to do with their physical condition. If it seems obvious that data from the deceased population greatly enhances the accuracy of human life expectancy, we may want to apply the same principle to a population of pipes experiencing breaks.

In other words, for proper pipe failure statistical and forecasting analyses, all pipes active during the whole or part of the period of breaks observation should be taken into account, including pipes that are no longer active.

Does your failure forecasting approach include abandoned pipes?

No abandoned pipe data? What should you do?

Missing abandoned pipes may not have been an issue if only a small percentage of pipes have been abandoned so far. But the relevance of that analytical requirement will only increase as utilities embark on more pipe replacements. Therefore, you should start collecting the information pertaining to abandoned pipes and their breaks.

It does present some challenges that can be addressed with the right framework. These points will be covered in **Article #5** ADD LINK to PLACE HOLDER? (article coming) that will focus on data issues. If need be, abandoned pipes that are missing can be recreated.

Analytical approaches that do not allow incorporating abandoned pipes (such as desktop scoring) should be phased out. It is one of the reasons why such approaches have a poor break prediction capacity (as seen in **Article #2** ADD LINK). If machine learning or advanced analytics are adopted, the model developed should take abandoned pipes into account.

Discover infraPLAN

infraPLAN can help you achieve savings on R&R expenses and still meet your LOF and other R&R objectives. We can show you how your municipality, large or small, can meet those goals.

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

About infraPLAN