PREDICTING PIPE BREAKS:

Desktop Scoring, Advanced Statistics (LEYP), and Machine Learning

Annie Vanrenterghem Raven and Kevin V. Campanella

Key Takeaways

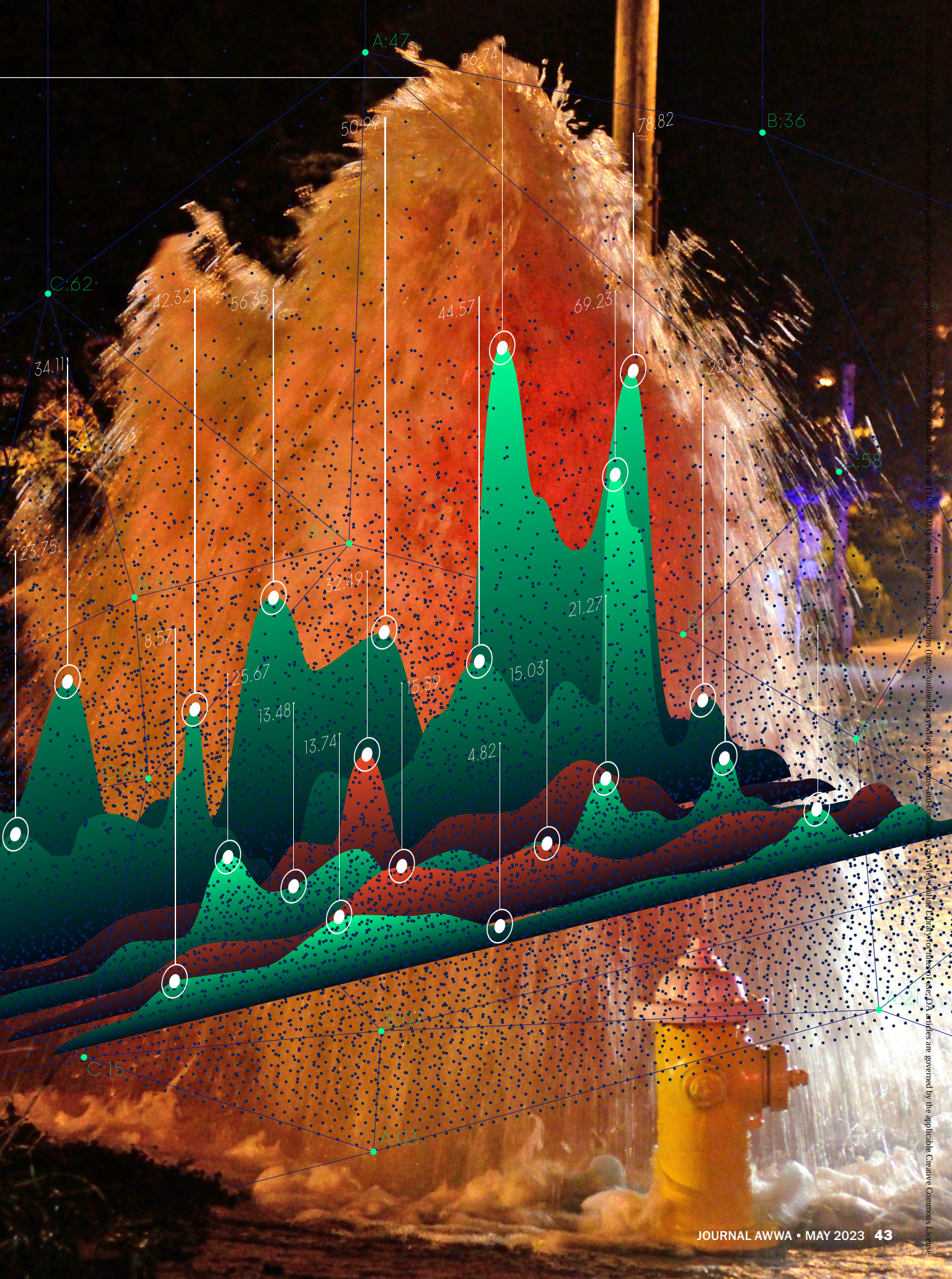
Using desktop scoring to determine the likelihood of failure of water pipes should be phased out and replaced with advanced analytics, particularly machine learning.

More accurate break predictions will lead to better estimates of how much to spend on pipe replacement and which pipes to replace.

Utilities should integrate information about abandoned pipes and breaks into failure forecasting, including machine learning models.

Some data issues can be cost-efficiently rectified using machine learning and automated algorithms.

Layout imagery by Maxger, TFoxFoto /Shutterstock.com

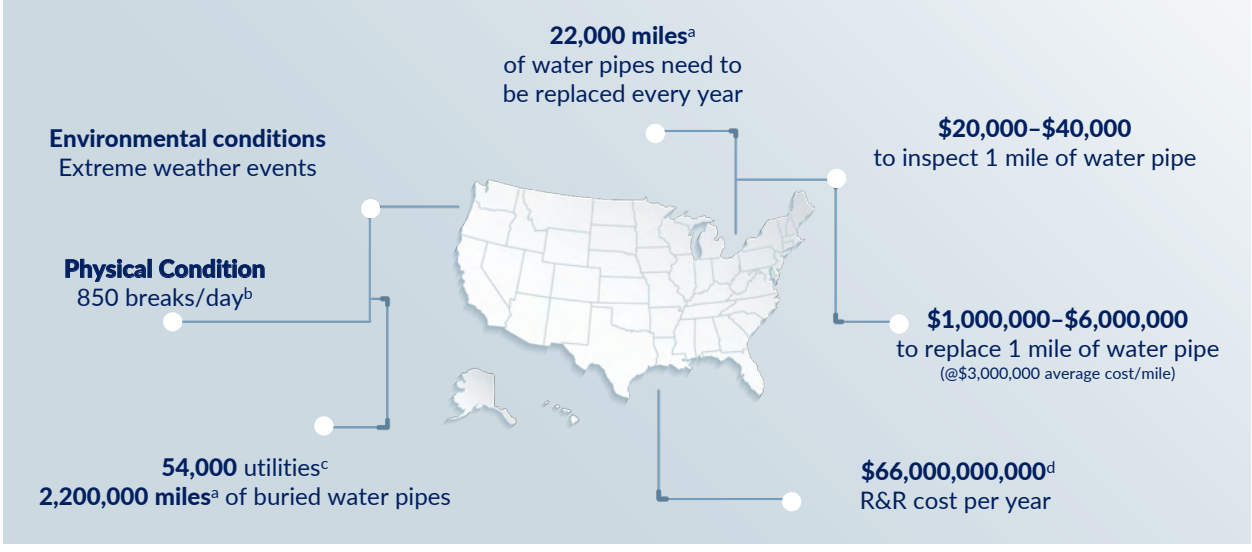


Maintaining the physical integrity of drinking water systems and maintaining records to support decisions about rehabilitation and replacement (R&R) require continuous attention and significant resources. Deferring R&R or making inappropriate R&R decisions can lead to enormous costs down the road and/or an abundance of costly breaks (and associated interruptions in service). In 2020, according to the Report Card for America's Infrastructure, published by the American Society of Civil Engineers (ASCE), the R&R of US water pipes (transmission and distribution) was expected to cost approximately $66 billion per year. Figure 1 shows the key contributions to the overall R&R needs for water pipes in the United States. Since 2020, costs have increased substantially as a result of supply chain constraints and rising inflation. Furthermore, the country has experienced extreme weather events that also have led to premature failures and additional R&R expenses.

Taking appropriate risk-based R&R actions results in better-scheduled pipe replacements, significant cost savings, and more reliable service. This approach requires that the likelihood of failure (LOF) of each pipe be assessed. There are several alternatives to generate that assessment.

For more accurate results, a pipe coupon can be analyzed in a laboratory, or the pipe can be inspected using a noninvasive technology. However, those two options come at a cost proportional to the number of coupons analyzed or miles of pipe inspected. At $30,000 per mile, inspecting the length of pipes slated every year for replacement (22,000 miles) would amount to $660 million per year. Alternatively, analyzing a utility's existing pipe and break data is a cost-effective option as it doesn't depend on the amount of data or miles of pipes. More amounts of useful data make analytics more accurate. It can serve as a screening step that identifies which pipes should be inspected in view of validating a decision to replace.

Rehabilitation and Replacement (R&R) of US Water Pipes in 2020



[a]Source: American Society of Civil Engineers, Report Card for America's Infrastructure 2020; 1% replacement rate equates to 22,000 miles of R&R per year.

[b]Source: Uni-Bell PVC Pipe Association, Water Main Break Clock: https://watermainbreakclock.com

[c]Source: US Environmental Protection Agency, 2008: *Factoids: Drinking Water and Ground Water Statistics for 2007.*

[d]22,000 miles of R&R per year × $3,000,000 (average cost of R&R for 1 mile of pipe) = $66,000,000,000 cost of R&R per year.

Figure 1

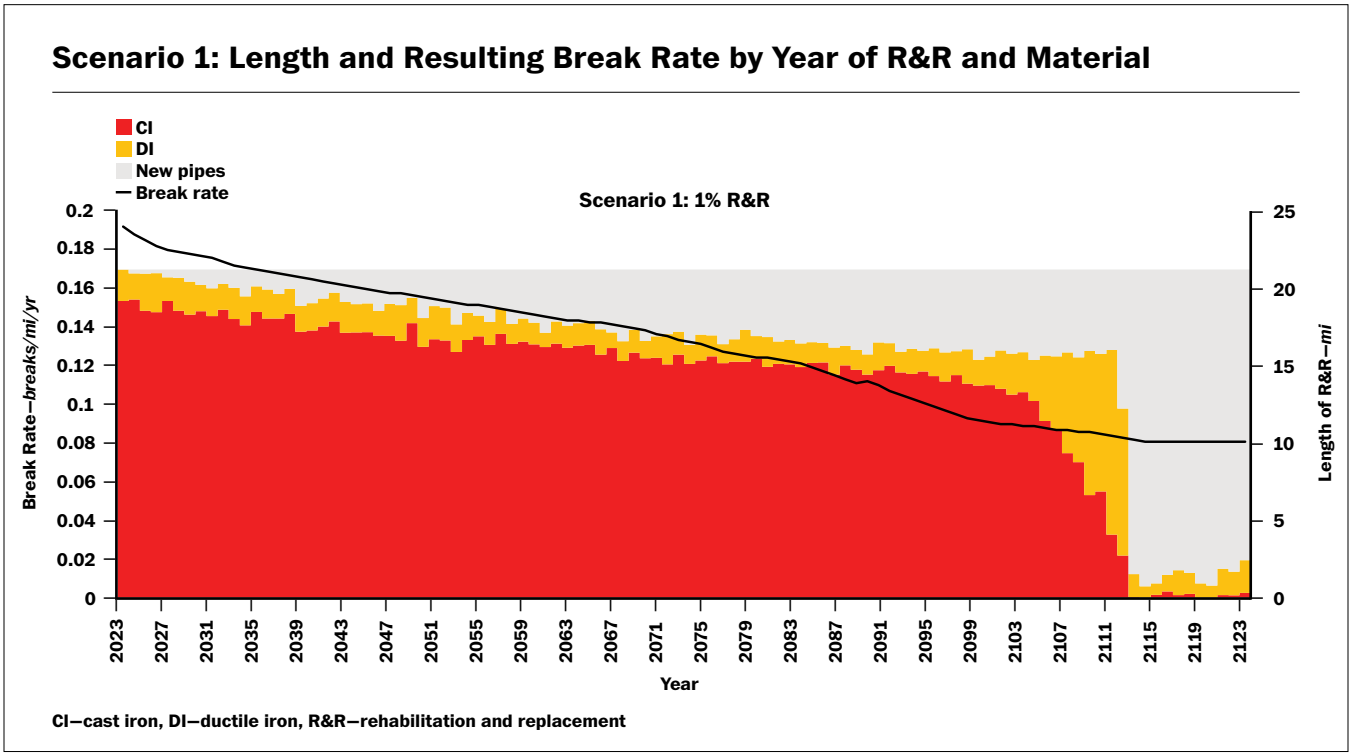


Figure 2

Data analysis can include spatial or desktop scoring using expert knowledge as well as advanced analytical approaches that include multivariable regression, or machine learning. Desktop scoring, a form of multicriteria analysis that typically results in assigning an LOF score of 1 to 5, is still frequently used. But if utility-specific pipe and break data are available at the pipe level, as is now common, advanced analytics will yield far more accurate and granular LOF results.

If it can respond to the specificity of pipe and break data, machine learning is particularly attractive because of its superior failure predictive capacity, analytical flexibility, and simplicity of use. Understanding machine learning, and therefore its adoption, is still limited in the water industry, which traditionally relies on tangible and field-based technical knowledge. Many water professionals are not yet comfortable with black box analytics, doubt the quality of their pipe and break data (including the ability to improve it at low cost), and may be reluctant to change long-held practices. To address these concerns, this article compares traditional desktop scoring with advanced analytics and presents limitations, capacities, and data issues and requirements, as well as remedies to improve data quality. We also list features prospective users should look for when considering machine learning to assess the quality of a break prediction model. In the next section, we first show the type of R&R plan optimization that can be obtained only with LOF scores generated through advanced analytics.

Machine learning is particularly attractive because of its superior failure predictive capacity, analytical flexibility, and simplicity of use.

Benefits of Advanced Analytics for Long-Term Planning

Advanced analytics—meaning approaches that go further than simply assigning LOF scores to pipes—allow utilities to improve their R&R planning capability, accuracy, and reliability while saving time and money. The following case study highlights a northeastern utility, with close to 2,100 miles of pipes (two-thirds cast iron and one-third ductile iron) with an average age of 57 years and a break rate of 0.200 breaks/mile/year.

If the utility's capital improvement plan (CIP) imposed an R&R rate of 1% per year (or 21 miles), this assumes its pipes would stay in service on average for 100 years. A machine learning–powered break forecasting model was used to estimate the LOF of each pipe for each year in the future. The future break rate in the absence of R&R was also projected to be 0.273 breaks/mile/year in 10 years, when the average pipe age would be 67, and 0.656 breaks/mile/year by 2065, when the average pipe would be 99 years old.

The resulting break rate after applying a 1% R&R rate at a cost of $63 million/year (21 miles × $3 million/mile), or $2.709 billion by 2065, was then estimated. R&R was made following two rules:

- Seventy percent of the annual R&R length includes pipes with the worst LOF, while 30% of the R&R length occurs for reasons other than physical condition.
- Replacement pipes are accounted for when estimating the resulting break rate; they are assumed to be more resilient than the existing pipes, with a break rate of 0.200 breaks/mile/year by the age of 85.

Figure 2 shows the break rate and the length of pipes replaced over time. Specific pipes are identified for each year (not just length to be replaced).

Systematically replacing 21 miles a year (1%) rapidly reduces the break rate beyond the utility's objective of maintaining the current break rate around 0.200 breaks/mile/year. This means that pipe replacement could be less aggressive while saving resources.

Scenario 2, shown in Figure 3, was proposed as an alternative to the replacement plan in Figure 2. In this case, the utility achieves its goal starting at 7 miles of annual pipeline R&R (the current level), ramps up to 19 miles by 2045, and then drops to around 10 miles of R&R per year by 2065. The total cost with this approach is $1.767 billion, which saves approximately $942 million (35%) compared with Scenario 1. This kind of CIP optimization cannot be conducted with LOF scores generated through desktop scoring.

Desktop Scoring

Desktop scoring requires that the variables that define a pipe, its environment, and anything that could speed up pipe degradation be first identified. Important information is needed:

- Pipe material, diameter, length
- Date of installation/abandonment
- Break history (number, date, cause)
- Soil conditions
- Operational conditions (e.g., pressure, anti-corrosion measures)
- Local conditions like traffic, groundwater, stray current, proximity to other infrastructure

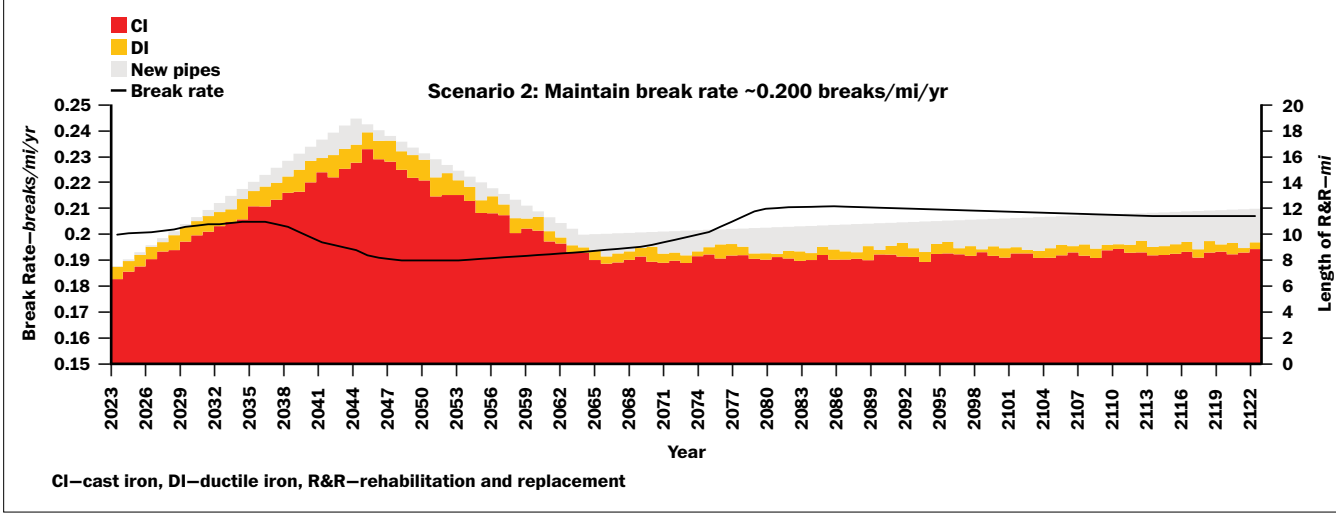


Figure 3

Furthermore, desktop scoring requires assigning values to these and other variables; each value is given a corresponding score (s). A weight (w) is finally assigned to each variable proportionally to fully characterize any degradation.

For each pipe, the score that corresponds to the value j of a variable i (s_{ij}) is multiplied by the weight that variable carries (w_i), and the LOF score of that pipe is the sum of the ($s_{ij} \times w_i$) entities for all the variables.

For example, take variables "material" and "soil." Material can take two possible values: cast iron (CI, score of 2) or ductile iron (DI, score of 1); the weight is 1. Soil can take the values "good" (score of 1) or "bad" (score of 2); the weight is 2. The LOF for those two pipes is

- pipe 1 (DI, bad soil); LOF = $1 \times 1 + 2 \times 2 = 5$ and
- pipe 2 (CI, good soil); LOF = $1 \times 2 + 2 \times 1 = 4$.

Characterization and scoring are typically done in Excel spreadsheets, using data from geographic information system (GIS) shape files, or directly in GIS.

The fundamental problem with desktop scoring is its subjectivity in assigning weights and scores, which can lead to LOF scores that are inaccurate. In the example, is a CI pipe ($s = 2$) two times more likely to break than a DI pipe ($s = 1$)? Does soil ($w = 2$) really contribute to pipe degradation twice as much as material ($w = 1$)? Is a DI pipe in bad soil (LOF score of 5) 25% more likely to break than a CI pipe in good soil (LOF score of 4)?

The interconnection between weights and values for different variables is another important consideration that cannot be addressed easily with desktop scoring. For example, the weight of soil may not be the same for CI or plastic, or a period of installation of 1970–1980 may yield different scores for two different materials (poor quality for material 1 at that time, but good quality for material 2). As a result, different scoring equations/models may be required for each value of certain variables, which makes it difficult to develop accurate relative LOF scores. A pipe with an overall score of 2 based on one equation could actually be in worse condition than a pipe with a score of 3 that was evaluated using another equation.

Regardless of how well they understand their system, it can be difficult for users to decide which variables should be granted their own model and equation, and how the output results from those equations compare. Doing this correctly requires careful statistical techniques that cannot be applied manually. Advanced analytics are the tools of choice that automatically tackle those scoring challenges.

Scoring presents an advantage: it allows users to estimate LOF scores when break data are not available at the pipe level. In these cases, there is no other option but to rely on expert opinions and experience. However, if each break is properly assigned to the pipe on which it occurred, the LOF of all the pipes can be more precisely predicted using advanced analytics. The next section compares advanced analytics and desktop scoring.

Advanced Analytics

Two advanced analytical approaches are described here: machine learning and the multivariable model, linear extended Yule process (LEYP). LEYP was developed by IRSTEA (National Research Institute of Science and Technology for Environment and Agriculture), based in France. The variables identified previously as required for desktop scoring are also needed for advanced statistics, with the only difference being that each break must be specifically assigned to the pipe on which it occurred. Both approaches use built-in statistical techniques that automatically identify the right weight and scores of each variable on the basis of the weights and scores of other variables. They then determine how pipes compare, alleviating the modeling difficulties encountered with desktop scoring described in the previous section. This makes assigning LOF effortless and more accurate for the user once the model has been properly configured by the data scientist.

Scoring presents an advantage: it allows users to estimate LOF scores when break data are not available at the pipe level.

Machine learning is superior to LEYP not only because it can make slightly better predictions but also because it doesn't need calibration by the user at each run. While a strong machine learning model must account for the specificity of pipe and break data, this is embedded in the machine learning software; a proper machine learning model is internally programmed to automatically make the kind of calibration choices that a user must make manually with LEYP for every model run. Specifically, LEYP calibration requires preliminary descriptive statistics to ensure that mathematical conditions are met for certain variables, especially quantitative continuous variables like year of installation and diameter.

Comparing Advanced Analytics and Desktop Scoring

Desktop scoring generates a relative LOF score (not a probability of failure) for each pipe currently (not in

the future). The result from LEYP or a machine learning model is a probability of failure translated into a predicted break number for each year in the future which, as seen in the case study, is valuable when simulating long-term R&R scenarios. Annual estimates consist of how many breaks will occur with various R&R investment levels (including "do nothing," current CIP, or any other replacement strategy). Seeing the effects on systemwide break rate and risk allows the utility to determine appropriate investments over a planning horizon. Desktop scoring doesn't provide this level of forecasting or guidance on appropriate long-term investment levels.

Future breaks are forecast by observing the past behavior of all the pipes that have experienced degradation, which includes pipes in service as well as those that have been abandoned. Properly configured advanced failure forecasting models allow accounting for data from abandoned pipes, which greatly improves their predictive capacity and accuracy, especially given that abandoned pipes tend to have experienced more breaks.

Regardless of the analytical approach, the quality of the results depends on the quality of the initial data and the model's calibration.

This is not an option with desktop scoring. While including abandoned pipes may have little consequence for systems that have not yet undertaken much R&R, the importance of abandoned pipes will only grow as the percentage of R&R increases.

Desktop scoring often ends up assigning the same simple 1–5 scores to many pipes, which makes prioritization of pipe inspections or replacements difficult and potentially inaccurate. LOF scores from advanced analytical approaches come as continuous digital values, probability of failure, or predicted break number, offering much richer granularity and analytical potential.

With desktop scoring or LEYP, the value of a variable must be known for every single pipe. For example, soil testing results cannot be introduced unless they are available for each pipe, which is rarely the case. In comparison, machine learning can draw inferences from incomplete information.

Models developed with data coming solely from the utility being analyzed tend to yield more accurate predictions (provided the data are available in sufficient quantity and quality) than "big models," where data from other utilities are added. However, if that utility's data quality is poor or insufficient, machine learning can (theoretically) "borrow" data from other utilities as a temporary measure while data from the utility being analyzed improves. This is not possible with desktop scoring or LEYP. The "big model" does require that a value attributed to a variable at one utility represents the same thing at another utility. Finally, as described in the next section, predictions of future breaks are far superior with advanced analytics, compared with desktop scoring.

Model Validation

Whether utilities can trust a model's predictions rests on validating model outputs with actual breaks. Examples in the following sections show validation results for two systems of different size and condition for two analytical approaches (desktop scoring, LEYP, and/or machine learning).

System 1: Large System With Medium Break Rate

1,490 Miles of CI Pipes and 714 Miles of DI Pipes

This study took place in 2022, with break data from 2002–2021; 2002–2019 breaks were used to generate a predicted LOF score with desktop scoring and machine learning for each pipe in 2020. The pipes were ranked from highest to lowest LOF. We then calculated the percentage of the total 2020–2021 actual breaks that would have been avoided had 1%, 5%, or 10% of those pipes (worst LOF first) been replaced by 2020. The higher the percentage of breaks avoided during the validation period, the more accurate the model.

Table 1 shows the validation results obtained with desktop scoring and machine learning for System 1. As shown in Table 1, if the top 5% of worst pipes, as ranked by their desktop scoring LOF, had been replaced, 19.4% of the 2020–2021 breaks would have been avoided. However, if ranked with machine learning using the 2020 predicted break number (PBN), that percentage becomes 79.3%. For the top 10% of worst pipes, the percentage avoided with desktop scoring was 45.2%, while for machine learning it was 84.9%.

These results indicate that the performance of the machine learning model was excellent, and it was almost two to four times better than with desktop scoring (depending on the percentage of pipes targeted for replacement).

System 2: Very Large System With High Break Rate

2,457.7 Miles of CI Pipes and 2,511.7 Miles of DI Pipes

This study took place in 2021, with 2005–2020 break

data. 2005–2018 breaks were used to predict LOF scores for 2019, and these results were compared with the actual 2019–2020 breaks.

Table 2 shows the validation results obtained with desktop scoring, LEYP, and machine learning. In this case, a LEYP model has to be created for each material, so to make valid comparisons, the desktop scoring and machine learning models were also validated separately for each material.

Results in Table 2 show that advanced analytics yield much better results than desktop scoring. The scale of the improvement depends on the percentage of pipes replaced, the material, and the approach. For this water system, CI pipes constitute a cohort for which predictions are more difficult to make regardless of the approach, indicating that the data need further improvement. Even for that rather weak cohort, results with advanced analytics are more than six times better at 5% replacement, and five times better at 10% replacement, than with desktop scoring.

These examples show that advanced analytics have a much higher failure forecasting capacity than desktop scoring, which is to be expected given the limitations of the desktop approach. Machine learning performed slightly better than LEYP, but the calibration requirements are much lighter. Machine learning is therefore a better option.

System 1: Model Validation Based on Breaks Avoided by All Pipes Replaced

All pipes *% replaced*	Model Type: % Breaks Avoided (Worst LOF First), 2020–2021		
	Desktop Scoring	LEYP	Machine Learning
1	17.8	NA	30.6
5	19.4	NA	79.3
10	45.2	NA	84.9

LEYP—linear extended Yule process, LOF—likelihood of failure

Table 1

System 2: Model Validation Based on Breaks Avoided by DI/CI Pipes Replaced

All pipes *% replaced*	Model Type: % Breaks Avoided (Worst LOF First), 2019–2020		
	Desktop Scoring	LEYP	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

CI—cast iron, DI—ductile iron, LEYP—linear extended Yule process, LOF—likelihood of failure

The cells under each model type contain the percentage of breaks avoided for (1) DI pipes and (2) CI pipes.

Table 2

Data Issues and Remedies

Regardless of the analytical approach, the quality of failure forecasting results depends on the quality of the initial data and the model's calibration. As described in the following sections, data issues may be the result of incidental human errors or faulty data recording and management processes.

Incidental Data Issues

Incidental data issues include missing, illogical, and incoherent values. These issues tend to be isolated, resulting mainly from incidental human errors or lack of consistent record keeping and data management. Because these issues are not the result of a faulty process, once they are corrected, the issues should be gone.

Identifying these issues is straightforward for missing or illogical values; for example, a date of break or of abandonment cannot be prior to a date of installation. In addition, analyzing data trends helps identify outliers. For example, as seen in Figure 4, which shows the number and length of pipes based on the year of installation at System 2, most DI pipes were installed after 1976. The utility flagged DI pipes with recorded year of installation prior to that year (some as early as 1915). Issues such as these can be quickly identified using an automated tool with an extensive library of known issues.

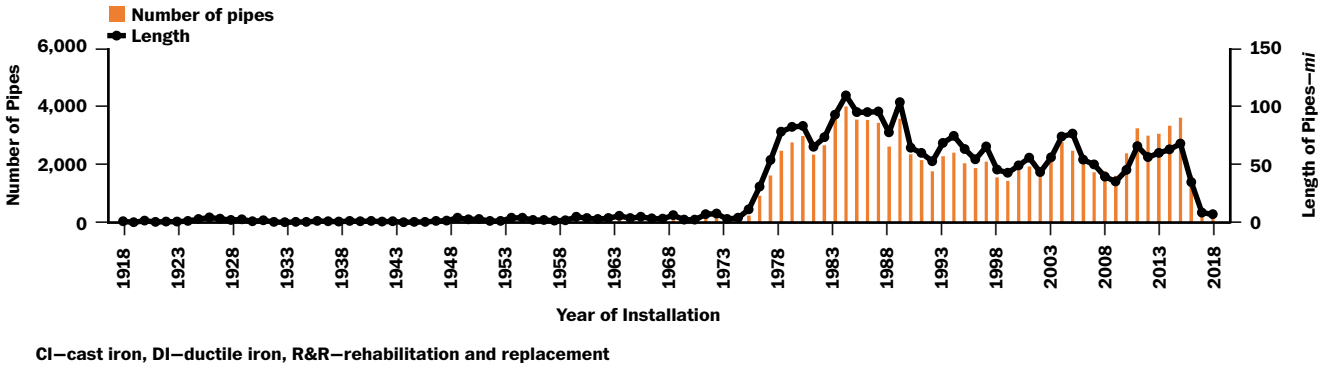


Figure 4

Structural Data Issues

Unlike incidental problems, structural data issues can be more complex, and if they are not fixed, errors will continue to surface even after existing issues have been cleaned up. Three main processes lead to structural issues: (1) failure to properly record abandoned pipes, (2) incorrectly linking breaks and pipes, and (3) failure to report in a computerized maintenance management system (CMMS) changes of pipe identifications (IDs) occurring in GIS. Because of these structural issues, breaks may end up lost to the study or assigned to the wrong pipe.

Unlike incidental problems, structural data issues can be more complex, and if they are not fixed, errors will continue to surface.

The absence of abandoned pipe records or the improper management of IDs at the time a pipe or a portion of a pipe is replaced results in the loss to the study of the breaks that occurred on those pipes or their being assigned to the wrong pipe. If a utility has not undertaken much R&R, these issues may not have had much effect, but they will become more prominent as the rate of R&R increases.

If breaks were previously recorded in spreadsheets or paper reports, their location may be in the form of the nearby address, and breaks are later "geocoded" in GIS by associating them with nearby pipe IDs. Because this process is often automated, wrong assignments can be made, especially if many pipes are in the area or if the break occurred on a pipe that was later abandoned and wasn't recorded.

Pipe and break data needed for statistical studies are collected mostly in GIS or CMMS not designed for such statistical analyses. For example, a CMMS is a work order depository that tracks work orders—i.e., who worked on what, when, for how long, and at what cost. While a break may be associated in the CMMS with the ID of the broken pipe at the time of the break, the pipe ID often is not updated in the CMMS if it evolves in GIS.

The above issues are illustrated with the following example. Pipe 105 is 200 feet long. It experienced a break in 2010 that was recorded in the CMMS at that time. Imagine that, in 2015, in the GIS, the ID of that pipe (105) became 106 as the result of a change in endpoints (no physical change). However, the break remained associated with pipe 105 in the CMMS, but that ID no longer exists in GIS. Therefore, the break is lost to the study. Or, in 2015, 100 feet of pipe 105 (the section that had a break in 2010) was replaced with a new material; that 100-foot section of pipe is a new GIS object that was given the recycled ID of 105; the remaining 100 feet was assigned a new ID, 106. The break remains associated in the CMMS with pipe 105, which exists in the GIS but is not the actual pipe that experienced the break. The break is assigned to the wrong pipe.

Resolving Data Issues

While incidental issues can be corrected one at a time by looking at the map or consulting source documents (if available), this is time-consuming, especially for systems that have a high percentage of data issues. To save time and resources, automated remedies based on statistical and spatial considerations using machine learning may be considered.

For example, returning to System 1, 2% of the length of pipe was missing material information, 4.5% did not have a year of installation, and 0.02% was not assigned a diameter. A machine learning data-cleaning module was used to determine those missing values. Model validation relied on picking 20% of the pipes that were not missing any information, deliberately removing some of those values, then using the cleaning module to estimate what they actually were. The material was properly predicted for 98.3% of the sample; the average difference (absolute value) between actual and predicted values was 1.08 years for the year of installation and 0.59 inches for the diameter.

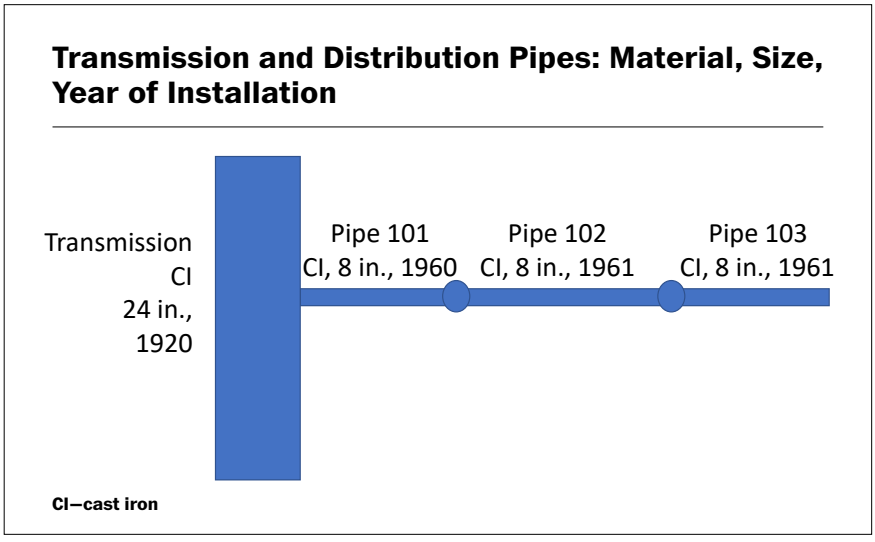


Figure 5

Machine learning models require less calibration effort from the user and lead to less human error.

However, not all data cleaning consists of filling in missing values. Situations are often more complex and require that more advanced cleaning algorithms be developed after observing the data for error patterns. Abandoned pipes may need to be retrieved, but this cannot be automated in GIS. The most complex variable to flag as incoherent and ultimately restore is the year of installation because it must follow a certain logic:

- Distribution pipes must typically emanate from a transmission pipe (unless there is a well).
- Along the path from any distribution pipe to its transmission line, the year of installation cannot increase (unless replacement occurred).

For example, in Figure 5, pipe 102 could not have been installed in 1965 if pipe 103 was installed in 1961. Such logic does not apply if replacement has occurred.

An algorithm ("pipedate") was developed to identify the most likely path of any distribution pipe to a transmission pipe, flag incoherence, and suggest coherent years of installation. It also identifies pipes that are likely replacement pipes and provides recommendations for the year of installation of the original abandoned pipe if that is missing.

Model Validation: Addressing Utility Concerns

As they consider incorporating machine learning into their failure forecasting and physical condition assessment programs, utility managers may want to consider several aspects pertaining to validation if they are to build trust in the results.

- Has the machine learning model been compared with at least one other advanced statistical model? In this case, a simple regression model based solely on age does not constitute a valid basis for comparison, as it should be easily outperformed by machine learning given that age is not a significant failure forecasting factor. Comparison with multivariable LEYP model results is a better basis for validation.
- If the utility has its own LOF scores (from desktop scoring, for example) and breaks are assigned to pipes, can a simple validation technique (similar to the one described previously that can be easily run in Excel) be used to evaluate the scoring model's predictive capacity?

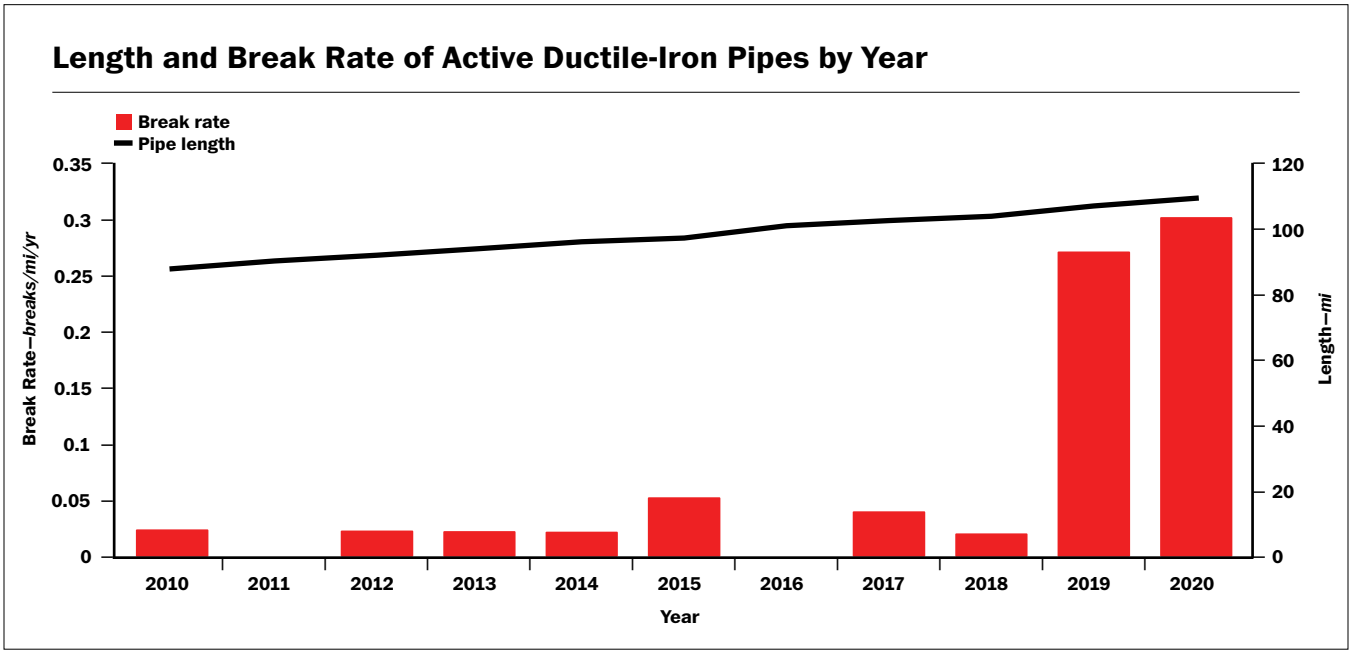


Figure 6

- Is the validation meaningful? It is recommended that the past break rate be evaluated before validation because it may be meaningless if the break rate during the test period is off trend. Figure 6 shows statistical results for a small system with approximately 100 miles of DI pipes. The break rate spiked in 2019–2020 as a result of operational changes. A validation model focusing on those years provides excellent (but misleading) validation results given that there were many more breaks during those two years.
- Are the data used in the model representative? As mentioned previously, a model's results are only as

Output Results From Machine Learning, 2021 PBN

Pipe ID	Length *ft*	Diameter *in.*	Material	Date Installed	2021 PBN	Breaks
201	888.8	8	CI	10/11/1972	0.0787	0
202	886.9	8	DI	1/1/1972	0.0318	0
203	37.3	8	CI	1/1/1915	0.0022	0
204	233.0	8	CI	1/1/1915	0.0208	0
205	96.2	8	CI	1/1/1938	0.2326	2
206	96.0	8	CI	1/1/1938	0.0069	0

CI—cast iron, DI—ductile iron, ID—identification, PBN—predicted break number

Table 3

Advanced analytics do not require any subjective choices; they yield excellent predictive results using utility-specific pipe and break data that in general are readily available or easily obtained.

good as the initial data. Utilities must ask themselves how trustworthy results are if they are based on, for example, 80% of the pipes and 50% of the breaks because the remaining data have various issues. Before any modeling project, the initial data should be thoroughly reviewed, and any issues should be identified and corrected, potentially taking advantage of automated processes to do this.

- Do results make sense when tested against simple descriptive statistics? Machine learning allows for multiple connections the human brain cannot follow, but the black box results should make sense when examined through the lens of simple descriptive statistics.
- To illustrate this point, we ran the following exercise with results from System 2: we compared the model output results (PBNs) of a series of two pipes with similar values for all variables but one, focusing on material for pipes 201 and 202 (they are both about 867 feet long, 8 inches in diameter, installed in 1972, with no previous breaks; one is CI, one DI), length for pipes 203 and 204, and number of breaks for pipes 205 and 206. This information is shown in Table 3, along with the PBN of each pipe for 2021.

Intuitively, we would expect a pipe that belongs to the group with a higher break rate, or that is longer, or has more historical breaks to have a higher PBN. Here, for example the PBN of CI pipe 201 (0.0787) is 2.5 times the PBN of DI pipe 202 (0.0318). This is to be expected given that the break rate (yearly average) of CI 8-inch pipes installed in 1972 had been previously found to be 0.671 breaks/mile/year, 1.3 times the break rate of similar DI pipes (0.504). Pipe 204 is more than six times longer than pipe 203, while the PBN is 9.5 times larger. The pipe that broke twice (205) has a PBN that is 33 times the PBN of the pipe that has had no break (206); this is also to be expected as the number of previous breaks is the most important predictive break factor. PBN results generated by machine learning must make sense when tested against simple statistics.

The Future of Pipe Breaks Prediction

This article has shown that desktop scoring has significant limitations as a tool to predict future breaks. Unlike desktop scoring, advanced analytics do not require any subjective choices; they yield excellent predictive results using utility-specific pipe and break data that in general are readily available or easily obtained.

Machine learning models also require less calibration effort from the user and lead to less human error. They present the greatest opportunity to save utilities money by creating a proactive pipe replacement plan that identifies the right pipes to replace at the right time.

Regardless of the approach, data quality is essential. Any issues with data quality must be identified and corrected, and that effort can be substantially streamlined by using automated tools also powered by machine learning. The processes leading to collecting and managing pipe and break data, especially abandoned pipes, may need to be restructured. Ideally, CMMS tools will add new features to better serve high-level analytics, and they will enhance the role those tools play in R&R planning.

Simple but rigorous descriptive statistics help interpret and validate what are still often perceived as black box results.

About the Authors

Annie Vanrenterghem Raven is the CEO of infraPLAN, New York, N.Y.; avanraven@infraplan-llc.com.

Kevin V. Campanella is the asset management and utility planning director for Burgess & Niple Inc.; Columbus, Ohio. He currently serves as the chair of AWWA's Asset Management Committee.

https://doi.org/10.1002/awwa.2089

AWWA Resources

- The Value Proposition for Likelihood-of-Failure Modeling. Kahn C. 2021. *Journal AWWA.* 113:1:30. https://doi.org/10.1002/awwa.1649
- Optimizing Cluster Selections for the Replacement Planning of Water Distribution Systems. Chen TY-J, Man C, Daly CM. 2021. *AWWA Water Science.* 3:4:e1230. https://doi.org/10.1002/aws2.1230
- Machine Learning for Pipe Condition Assessments. Fitchett JC, Karadimitriou K, West Z, et al. 2020. *Journal AWWA.* 112:5:50. https://doi.org/10.1002/awwa.1501

These resources have been supplied by *Journal AWWA* staff. For information on these and other AWWA resources, visit www.awwa.org.