



Utilizing Predictive Modeling to Identify Lead Service Lines (LSLs): Rationale and Guidance Documents

May 2023

NJ Department of Environmental Protection

Division of Science and Research

Authors and affiliations:

Lori A. Lester, Ph.D., Chief, Bureau of Risk Analysis

What was the main purpose of these rationale and guidance documents?

The Bureau of Safe Drinking Water (BSDW) requested that the Division of Science and Research (DSR) assess whether predictive modeling is as effective as other methods to locate lead service lines. In response, DSR wrote a [rationale](#) that reviewed and summarized the available literature regarding the usage of predictive modeling to locate lead drinking water service lines (NJDEP 2022a). Following review of this rationale, BSDW further requested that DSR write a [guidance](#) document targeted towards water suppliers to provide more detailed information on the best statistical practices for utilizing predictive modeling to determine the likelihood of a property containing a lead service line (NJDEP 2022b).

Overall, what did the studies show?

Predictive modeling is an effective and scientifically-sound approach to determine potential lead service line locations. Although predictive modeling has its own limitations and uncertainties, in the majority of the studies considered, the uncertainty associated with other techniques (e.g., historical records, excavation, etc.) tended to be higher than the uncertainty associated with predictive modeling results. Furthermore, by replacing the lead service lines based on the more accurately predicted locations from models, the resources needed for lead service line replacement projects would be greatly reduced. Therefore, predictive modeling is an accurate and cost-effective approach to identify lead service line locations.

Would predictive modeling be a useful tool for my water system?

Predictive modeling may be a useful tool for some water suppliers. Suppliers are encouraged to consider various parameters (e.g., historical records accuracy, construction timing, water system size, and number of service lines of unknown material) to determine whether predictive modeling would be useful for their system. Modeling would likely be most useful for large water systems with inaccurate or incomplete records, with many buildings constructed prior to the 1986 amendment to the Safe Drinking Water Act to ban lead, and with many service lines of unknown material.

How can the results of a predictive model be utilized?

The results from the predictive models (i.e., the likelihood of a lead service line being present at each building) can be utilized in two ways: (1) to prioritize locations for service line replacement and/or (2) to support inventory development by determining which service lines may be counted as lead in the inventory and which service lines will remain unknown requiring additional investigation or physical verification to confirm the service line materials. Additional information regarding meeting inventory requirements can be found in the Department's [Guidance for Developing a Lead Service Line Replacement Plan](#) (NJDEP 2022c).

What is a “Predictive Modeling Report”? When should we submit this report to the Department?

If a water supplier decides to use predictive modeling to assist in the development of its inventory of current lead service lines, then it should consider submitting a “Predictive Modeling Report” annually to the Department along with its inventory. This report should include details about (1) how the model was developed, evaluated, and assessed for accuracy, (2) the results of the predictive model, and (3) how the results were used to inform prioritization of properties for replacement and/or develop an inventory. Details on how predictive modeling is used to inform service line replacements should also be incorporated in their annual Replacement Plan (NJDEP 2022c). See N.J.S.A. 58:12A-44.

How should data be organized?

The guidance document provides additional information and recommendations on how to develop a data management plan. Essentially, all the historical data associated with each property's service line should be collected. These data should be stored in a spreadsheet or database program (e.g., Microsoft Excel or Access). Each row in the dataset should be one property, and then the columns should contain the different data available. There should be separate columns for utility versus property owner side of the service line. Old physical records such as notecards should be digitized in the database. Finally, a data dictionary should be included that explains what data are contained in each column of the database.

How can we test whether our historical records are accurate?

Some historical records may be inaccurate which may result in lead service lines being found in locations where they are not expected or vice versa. To assess whether your system's historical records are accurate, track what materials are found as service lines are inspected and replaced. These results can then be compared to the historical records to see whether they are accurate. In the guidance document, we explain how to develop a “historical records materials confusion matrix” which may be a helpful tool to determine whether your historical records are accurate.

What data should we use to run a predictive model?

The Department endorses using a randomized sampling approach to select a subset of service lines to use as input to the predictive model. This means that each building serviced by the system should have an equal chance of being selected as a sample. Additional details are included in the guidance to describe how to decide what area of interest to study and how to determine the appropriate sample size. The service line material at each property in the sample should be physically verified. This sample can then be utilized to run a predictive model. The model should then be rerun annually using new data as replacements occur.

How do we ensure that our predictive model is accurate?

The accuracy of the model should be assessed during all stages of development and usage to ensure that the resulting predictions are accurate and appropriate to use for decision-making. In particular, the model accuracy must be verified in at least two ways: (1) by comparing the model results to hold-out data from the original randomized sample and (2) by comparing the results to field results during inspections and replacements.

How can we use the predictive modeling results to inform our inventory?

A water supplier may utilize modeling results to inform its inventory, but it should clearly explain to the Department (in their annual Predictive Modeling Report) how it determined which service lines were included as lead in their inventory. In the guidance document, the Department suggests two potential ways to make this determination: (1) the inflection point technique or (2) the threshold technique. For further information regarding inventory requirements, please see the Department's [Water Supply](#) and [Lead](#) websites.

Does the Department support the use of predictive modeling?

Yes. Predictive modeling may be useful for some water systems to inform decision-making, to plan strategically, and to protect human health. However, modeling is only one tool for decision-making, it is optional, and it is not required under a regulation or a statute. Thus, water suppliers are encouraged to consider other components such as equity, logistical constraints, and high-risk populations in their decision-making for prioritizing how to locate and identify lead service lines. In conclusion, the Department encourages suppliers to utilize predictive modeling, if deemed useful. The Department also reserves the right to reject the threshold, request physical verification of additional service lines, or request additional information about the model.

My water system does not plan to use predictive modeling as a tool. Should I still read the rationale and guidance documents?

Yes! The guidance document may still be useful, especially the sections on how to manage data and how to be transparent with the Department.

Who should be contacted for further questions?

Questions about predictive modeling can be sent to the Water Supply Inbox at watersupply@dep.nj.gov, and questions about financial assistance can be sent to waterbankinfo@dep.nj.gov.

References:

NJDEP. 2022a. [Rationale for Utilizing Predictive Modeling to Identify Lead Service Lines](#). (Ed. L.A. Lester) Trenton, NJ. 8 pp.

NJDEP. 2022b. [Guidance for Utilizing Predictive Modeling to Identify Lead Service Lines \(LSLs\) for Inventory Development](#). (Ed. L.A. Lester) Trenton, NJ. 9 pp.

NJDEP. 2022c. [Guidance for Developing a Lead Service Line Replacement Plan](#). Trenton, NJ. 8 pp.

State of New Jersey
Phil Murphy, Governor

Department of Environmental Protection
Shawn M. LaTourette, Commissioner



Division of Science & Research
Nicholas A. Procopio, Ph.D., Director

Visit the DSR website
<https://dep.nj.gov/dsr>