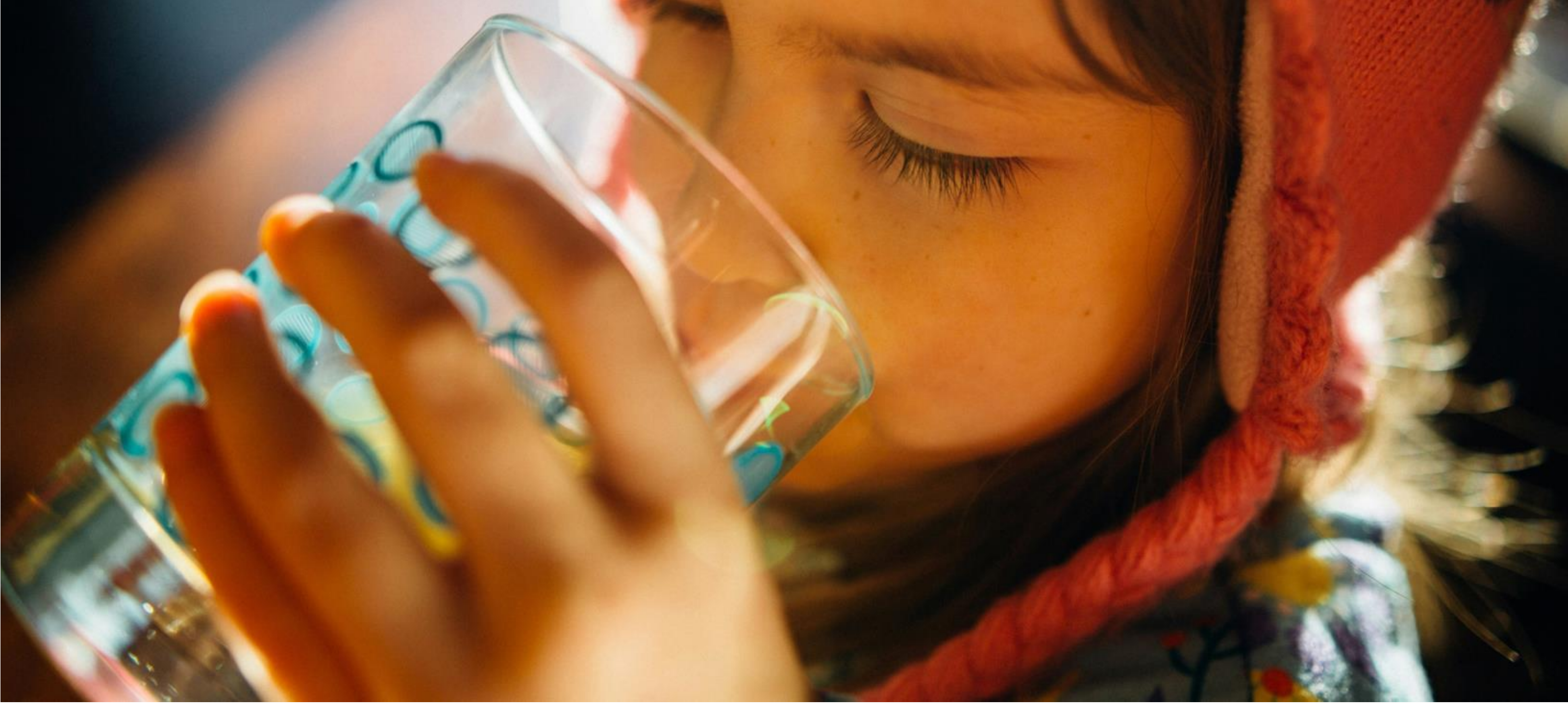


Development and Deployment of a Lead Service Line Inventory Application for North Carolina Water Utilities

Jordan Clark, Nicholas Institute for Energy, Environment & Sustainability, Duke University

Acknowledgments

We gratefully acknowledge the North Carolina Rural Water Association for their partnership and the utilities across North Carolina whose feedback was invaluable in developing this tool.

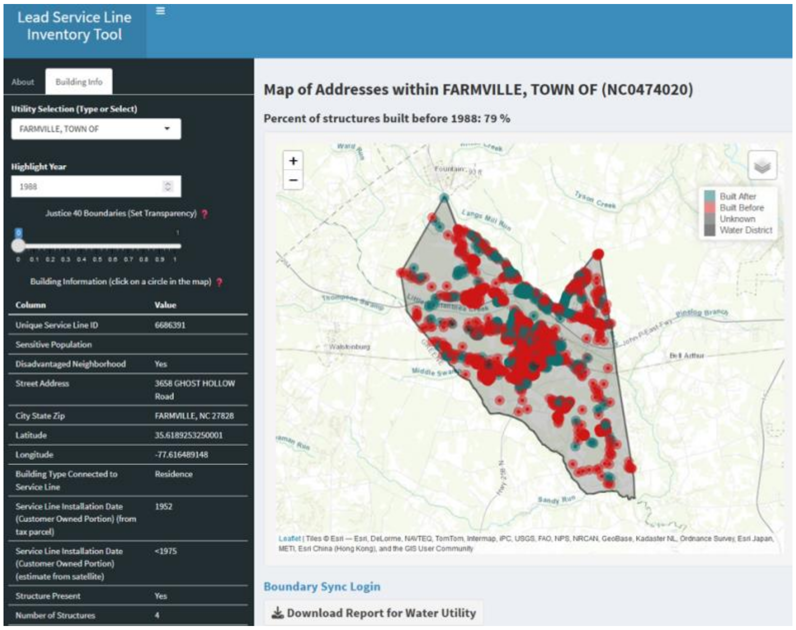
Funding for this work was provided by the BHP Foundation.

INTRODUCTION

In December 2021, the US Environmental Protection Agency (EPA) announced new Lead and Copper Rule Improvements requiring that community and noncommunity/nontransient water systems, regardless of size, develop an inventory identifying the potential presence of lead within each service line connection by October 16, 2024 (EPA 2021). This new mandate spotlighted a significant challenge for many water utilities, especially smaller and rural ones. The identification and replacement of lead service lines is critical given the severe health risks associated with lead consumption. However, the high costs of geographic information system (GIS) software and the technical skill required to effectively use such platforms have left many utilities at a disadvantage, unable to initiate or progress in their inventory efforts.

In response to this need, the Water Policy Program at Duke University's Nicholas Institute for Energy, Environment & Sustainability and the North Carolina Rural Water Association collaborated to develop a lead service line inventory application for North Carolina utilities (IoW 2024). This endeavor aims to not only facilitate compliance with EPA mandates but also to empower North Carolina utilities, regardless of size or resources, with the means to efficiently navigate the complexities of lead service line inventory development.

Figure 1. Visual overview of the Lead Service Line Inventory Tool



PROBLEM

The EPA's lead service line inventory rule places a significant workload on utilities. This task is particularly daunting for those with constrained resources. Many smaller utilities face challenges resulting from a dearth of comprehensive datasets created and maintained by the utility, lack of resources to use datasets like those employed here, absence of digital service area boundaries, and reliance on outdated paper records for older structures.

The federal government banned the use of lead in new plumbing systems in 1986, making a structure's construction year a key indicator of lead use in the customer-owned portion of service lines. However, an extensive review in North Carolina revealed that more than half of the tax parcel data either lacked year-built information or contained inaccuracies. Because an accurate year of construction for most structures in North Carolina was unavailable, utilities lacked access to the baseline data they needed to begin their inventories. As a result, many utilities did not know where to start.

SOLUTION

The solution is a user-friendly application designed for water utilities, which leverages publicly available data to provide a basis upon which utilities can more easily begin their lead service line inventories. This eliminates the need for utilities to provide their own data or invest in expensive GIS technology. The primary data sources integrated into this solution include the following:

- Utility digital service area boundaries
- Tax parcel data
- The National Address Database from the US Department of Transportation (DOT)
- The Global Human Settlement Layer (GHSL)

An underlying critical element to this work was a utility with a digital service area boundary to intersect with these data. While some utilities in North Carolina have digital service area boundaries, the majority (largely located in small, rural counties) do not. In addition to the use of the specific solution detailed here, utilities that did not have digital service area boundaries were encouraged to create them using a tool called BoundarySync (IoW n.d.). BoundarySync—created by the Internet of Water Coalition, which the Nicholas Institute leads in collaboration with partners at the Center for Geospatial Solutions at the Lincoln

Institute—is a free mapping platform that enables utilities to digitize their service area boundaries. Users can update and refine existing digital boundaries or digitize paper-based boundary images and files such as PDFs. This tool plays a pivotal role in bridging the technology gap by simplifying the creation and management of digital service area boundaries.

Since knowing the construction year of a structure is a strong indicator of whether the customer-owned portion is likely to contain lead, the initial solution planned to draw solely upon tax parcel data for this information. However, in North Carolina, more than 50% of these data either did not contain the year-built information or the included date was inaccurate.

To address these data gaps, we incorporated data from the GHSL, which is provided by European Commission Joint Research Centre. The GHSL project produces global spatial information about human presence derived from various data, including satellite imagery, census data, and volunteered geographic information. Specifically, we used GHSL data containing the (1) proportion of developed surface and (2) estimated building volume, a gridded dataset with a spatial resolution of 100 m.

Analyzing these variables over time allowed us to approximate the construction years of structures within each 100 m grid cell. The methodology used the values for 2020 (of both the proportion of the cells with development and the estimated building volume) as a reference point for analyzing the data retrospectively. To do this, we compared the proportion of increase in developed surface area and building volume at each five-year interval from 1975 to 2020. This analysis allowed us to determine or approximate the year of construction based on significant changes observed in these intervals. Minor changes between intervals were not considered indicative of construction activity in those years. Only increases representing more than a 10% change were deemed significant enough to suggest construction during a particular period. This approach enabled us to fill notable gaps in the data on the construction year of structures, especially in areas where traditional records were incomplete or inaccurate.

Throughout the development of this solution, we traveled to several training sessions hosted by the North Carolina Rural Water Association across the state in the fall of 2023. This allowed for direct interaction and engagement with representatives from 200+ water utilities, providing a powerful interface to demonstrate the tool and gain valuable feedback. Through these engagements, three additional elements were incorporated:

- Address data from DOT's National Address Database (DOT 2024) were included alongside the tax parcel data (NC OneMap n.d.). Sometimes there are multiple addresses within one parcel, and thus multiple water connections/meters, such as in multifamily dwellings and apartment complexes.
- In addition to showing the construction year estimated from the GHSL data, the construction year from the tax parcel data itself was also included (although in many cases this was an empty field). This approach significantly enhanced utilities' ability to transparently understand the process and data involved. It empowered them to make informed decisions on which data source to prioritize for a first estimate, leveraging their local knowledge for greater accuracy.
- Data from the Justice40 Initiative were incorporated, as one requirement of the inventory process is to identify if the structure is within a disadvantaged community (Biden 2021). This allowed for an initial identification of structures within disadvantaged communities. If the Justice40 data did not indicate that a structure was within a disadvantaged community, no information was specified in this field. By leaving this field blank, rather than indicating that a given structure was not in a disadvantaged community, we aim to encourage the utility to investigate further and confirm the status of a specific structure. To provide additional data required for the inventory, we also used tax parcel codes to discern potential sensitive subpopulations (e.g., day cares, schools, multifamily homes).

These engagements also informed the development of a web-based graphical user interface (UI) that allows users to dynamically view the data. Users select their utility from a dropdown menu to pull up a map that displays the utility's boundary, with dots representing each address, categorized by construction date relative to the lead ban. A key feature of the tool is its data export capability. The tool allows utilities to download a Microsoft Excel spreadsheet of information on each structure in their service area, creating a baseline more streamlined inventory management.

Not only does this allow the user to download the information specific to their utility's digital service area boundary, but it provides (1) clear and concise information that can be used within the inventory to meet the requirement to describe the data used and how they made the designation as lead or non-lead and (2) formatted data columns that largely align with North Carolina's inventory template (each state has their own template for this inventory.)

The iterative development process employed here, characterized by continuous feedback and refinement, has ensured that the solution is aligned with the practical requirements of its users. By engaging directly with utility representatives and incorporating their feedback, the project team fine-tuned the tool to meet their needs. Through this approach, the tool has evolved into an important asset for utilities to begin their inventory process.

CHALLENGES AND ADAPTATIONS

Throughout its development, the lead service line inventory tool faced several significant challenges, each necessitating adaptations and strategic response:

- **Data Accuracy and Urban Bias.** Discrepancies in data accuracy and completeness in the tax parcel data between urban and rural areas presented an obstacle. Urban areas typically had more detailed year-built records compared to rural regions. Thus, efforts to robustly validate the results provided by this solution were challenged with a large bias towards urban patterns of development.
- **Satellite Imagery Limitations.** The use of GHSL data, marked by five-year intervals between observations, presents challenges in accurately determining construction dates around the 1986 lead ban. With this interval, comparing construction activity or building volume between 1985 and 1990 means structures built between 1987 and 1990 would be characterized as having been built in 1985. Therefore, buildings constructed shortly after the ban would be categorized in our data as *preban* because of the timing of data collection. However, based on engagements with utilities, it was deemed appropriate to allow a one-to-two-year grace period after the lead ban as there may have been contractors with large amounts of lead-based materials on hand who may have used up their remaining supplies in the short time following the lead ban. Ultimately, the data interval challenge places constructions near the lead ban period in an uncertain category, but ultimately affects only a small portion of the dataset and is a narrow three-to-four-year time frame of uncertainty. Although the GHSL data provides crucial insights here, it was not originally intended for precise identification in this context. It acts as a foundational resource where no prior data exists, offering a significant step forward in identifying potential lead service lines despite its limitations.
- **State Variations in Inventory Requirements.** The potential for nationwide expansion brought to light the challenge of adapting to various state-specific inventory requirements. This necessitated a flexible and adaptable tool design, capable of accommodating these differences without compromising functionality or accuracy.

STRATEGIES FOR NATIONAL EXPANSION AND ENHANCED EFFECTIVENESS

If expanded, this tool could offer a path forward for utilities nationwide to efficiently tackle lead service line inventories. Drawing upon data available across the United States, the tool is well-positioned for national scaling However, certain key areas of additional work are crucial to enhance its effectiveness and adaptability at a national scale:

- **Integration with Private Well Data.** To refine the distinction between properties relying on public versus private water sources, incorporating private well data would enable a more accurate filtering process. This would ensure the inventory focuses on public water lines, which are pertinent for the lead service line inventory requirement.

- **On-the-Ground Validation.** Especially in rural areas where year-built data may be less reliable, conducting ground validation efforts would strengthen the accuracy of the tool's estimates. This would involve verifying the tool's data against actual structures in rural areas and, generally, across diverse geographical settings to ensure its reliability.
- **Refined (UI).** Enhancing the tool's UI to offer more engaging and intuitive data visualization would significantly improve the user experience. A more interactive interface could facilitate deeper insights and easier navigation for utilities in managing their lead service line inventories.
- **Incorporation of Socioeconomic and Health Data.** Adding layers of socioeconomic and public health data could offer utilities additional context for prioritizing lead line replacements. Understanding the demographic and health landscape of the areas served could guide more targeted and impactful interventions, especially in communities most vulnerable to lead exposure.

These enhancements would not only improve the tool's functionality and user experience but also ensure a comprehensive approach to identifying and addressing lead service lines. Such advancements would pave the way for a more informed, equitable, and efficient response to the public health challenge of lead in drinking water, making a crucial difference in communities nationwide.

CONCLUSION

The development and deployment of the lead service line inventory tool embody a significant step forward in leveraging technology to address public health challenges. Through a process marked by collaboration, innovation, and adaptation, this tool has emerged as a vital resource for utilities, especially those in under-resourced areas, to comply with the EPA's mandate. This initiative not only facilitates a more efficient and accurate inventory process but also reinforces the commitment to safeguarding public health through improved water infrastructure. The journey of the tool's development reflects a broader mission: to ensure that technological solutions meet the pressing needs of communities and regulators alike, making a safer environment for all.

REFERENCES

Biden, J. 2021. *Executive Order on Tackling the Climate Crisis at Home and Abroad*. Washington, DC: The White House. https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/executive-order-on-tackling-the-climate-crisis-at-home-and-abroad/.

DOT. 2024. *National Address Database*. Washington, DC: US Department of Transportation, Federal Highway Administration. https://www.transportation.gov/gis/national-address-database.

EPA. 2021. *Revised Lead and Copper Rule*. Washington, DC: US Environmental Protection Agency. https://www.epa.gov/ground-water-and-drinking-water/revised-lead-and-copper-rule.

European Commission Joint Research Centre. n.d. *Global Human Settlement Layer*. Brussels: European Commission. https://ghsl.jrc.ec.europa.eu/.

IoW. n.d. *Boundarysync* Cambridge, MA: Center for Geospatial Solutions, Lincoln Institute of Land Policy. https://internetofwater.org.

IoW. 2024. *Lead Service Line Inventory Tool*. Cambridge, MA: Center for Geospatial Solutions, Lincoln Institute of Land Policy. https://internetofwater.shinyapps.io/iow_lead_dash/.

NC OneMap. n.d. *NC Parcels*. Raleigh, NC: NC Geographic Information Coordinating Council. https://www.nconemap.gov/pages/parcels.

CITATION

Clark J. 2024. "Development and Deployment of a Lead Service Line Inventory Application for North Carolina Water Utilities." WP 01-24. Durham, NC: Nicholas Institute for Energy, Environment & Sustainability, Duke University. https://nicholasinstitute.duke.edu/publications/development-and-deployment-lead-service-line-inventory-application-north-carolina.

Header image courtesy Johnny McClung via Unsplash

Contact

Nicholas Institute | Duke University | P.O. Box 90467 | Durham, NC 27708
1201 Pennsylvania Avenue NW | Suite 500 | Washington, DC 20004
919.613.1305 | nicholasinstitute@duke.edu