Denver Water's LSL Inventory Predictive Model and LSLR Prioritization

Presented by Quirien Muylwyk, AECOM

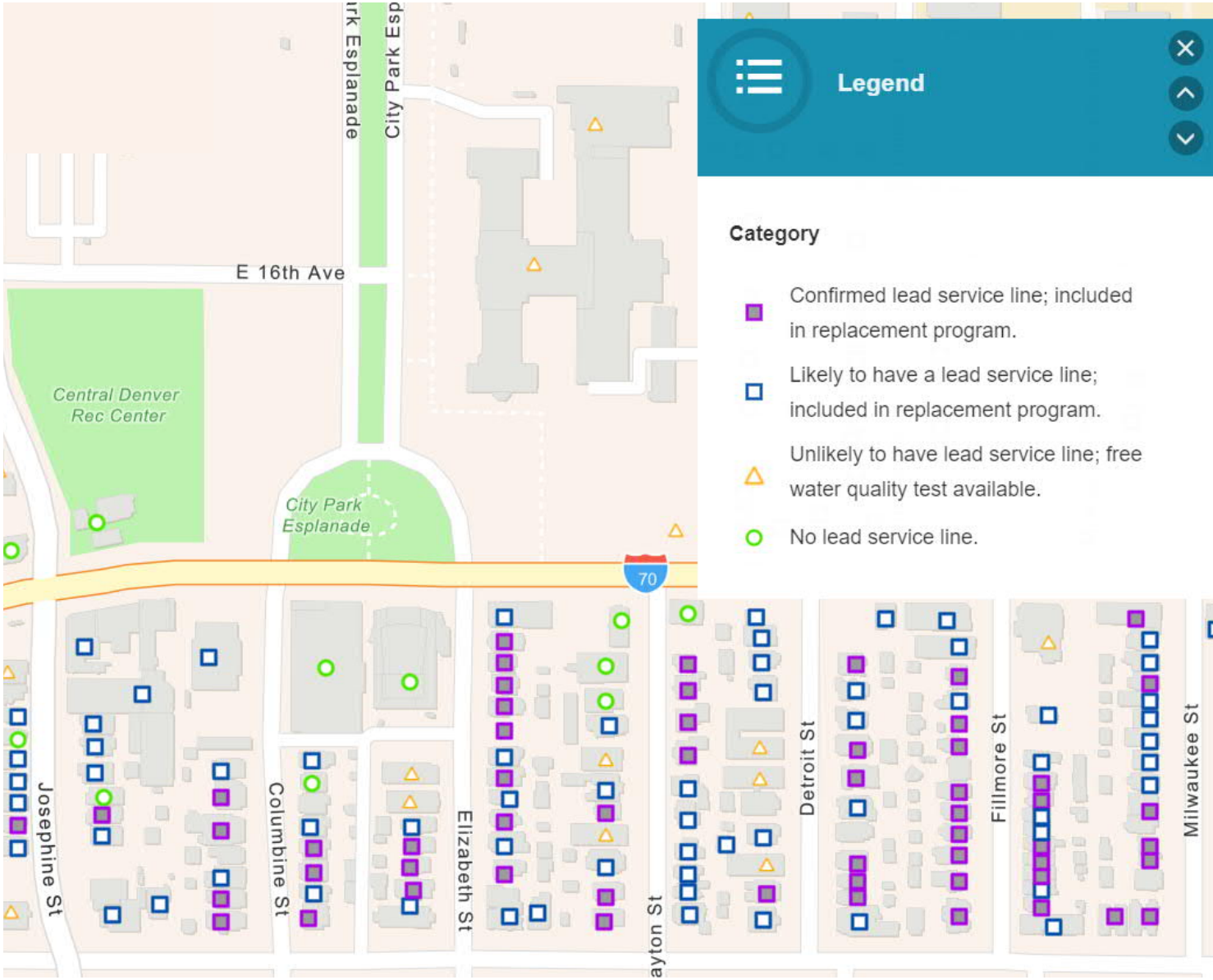
April 10, 2020 • ASDWA Lead Service Line Predictive Tools Webinar

Why use a predictive model to build Denver Water's LSL inventory?

We know how many LSLs there are likely to be in Denver Water's system based on statistics

But the statistics don't tell us where the LSLs are

This is an example of how we use the predictive model to help a water utility find the lead



Introduction to Denver Water

Founded in 1918 and now serves 1.4 million people

3,000 miles of pipe and approx. 350,000 services

There is no lead in water delivered to customers

Lead can get into water as it moves through customers' lead-containing household plumbing and service lines

pH adjustment used for corrosion control treatment

Denver Water's Lead Reduction Program Plan 2020

- All LRP tasks rely on the lead inventory and use it to target:
  - Outreach and education
  - Filter distribution
  - LSL replacement
- Service line is owned by the customer
- Historical records were poor

We needed to know the likelihood – and consequence – of finding lead at each property

Corrosion Control Treatment

- Implement pH/Alkalinity adjustment starting March 2020

Lead Service Line Inventory

- Continuous investigation to improve inventory
- Keep an up to date online map publicly available
- Estimated 64,000-84,000 LSLs

Accelerated Lead Service Line Replacement

- Begin construction in March 2020
- Target 4,477 LSL replacements
- Work areas within the City and County of Denver

Filter Program

- Begin initial distribution in March
- Complete distribution by August
- Continue filter replacement distribution until LSLs are replaced

Communications, outreach and education

- Incorporated into every program element

Overview of Denver's Accelerated LSL Replacement Program

- Denver Water estimates there are 64,000 – 84,000 LSLs
  - Remember, the customer owns the LSL
- The LRP targets approximately 5000 LSL replacements a year
- This means that it will take 15 years to replace all the lead services

With this number of LSLs — and with this goal for annual replacements — we needed a predictive model to prioritize our efforts to find and replace the lead.

Building a predictive model

What is a predictive model and how does it apply to lead inventories?

- Looking for patterns in the data to develop “rules” or an algorithm
- The predictive model “learns” from patterns how to assign a classification to the data ***(= will we find lead?)***
- The model is built using an initial data set ***(= initial inventory)***
- The model is trained using an additional data set ***(= field data)***
- The model is continually trained as more data are added ***(= more field data)***

We use the predictive model to efficiently build and update the inventory

We use the prioritization model to action and message the inventory

Context for building the predictive model

- Reduce the uncertainty in the inventory
- Build the inventory quickly
  - Less than four months to build the model before submitting the Request for Variance
- Because of the numbers involved, we used a predictive model to systematically check the rules and assumptions to refine the inventory
- Predictive model is the tool to focus the investigations used to improve the inventory as fast as possible

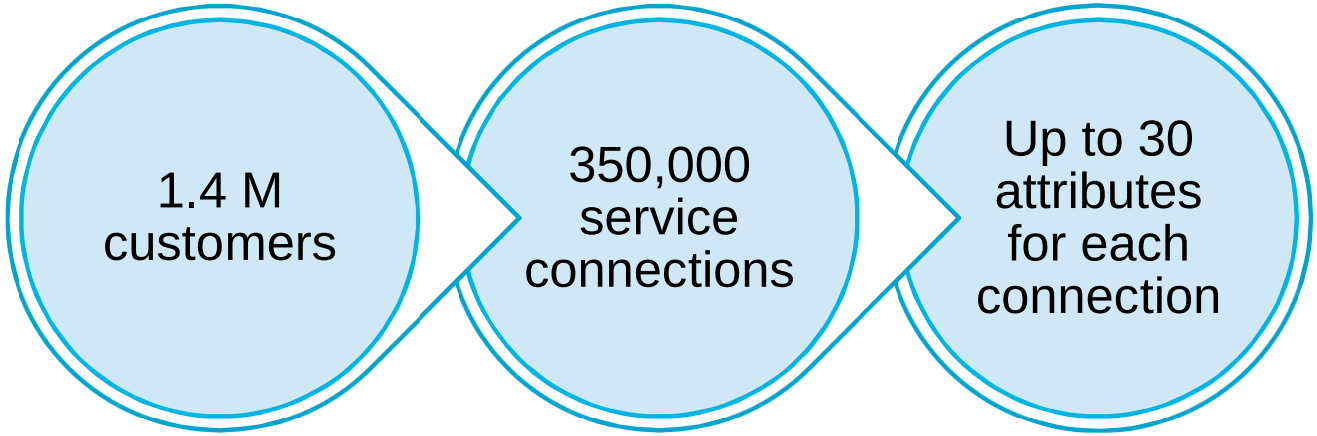
From lead inventory to predictive model

- A team is needed to assemble the tools and rules to set up the initial inventory
- Know the history of the system
  - Lead use
  - Partial replacements
  - How other assets and infrastructure are used

Examples of data used in initial inventory:

- ✓ **House age or build date**
- ✓ **Tap date**
- ✓ **Connection date to waterman**
- ✓ **Service line diameter**
- ✓ **Geographical trends**
- ✓ **Policies**

Lots and lots (and lots) of data . . .



The predictive model allows us to systematically review each property against the available data and to respond quickly to new rules and data assumptions.

How to train the predictive model

Train the predictive model

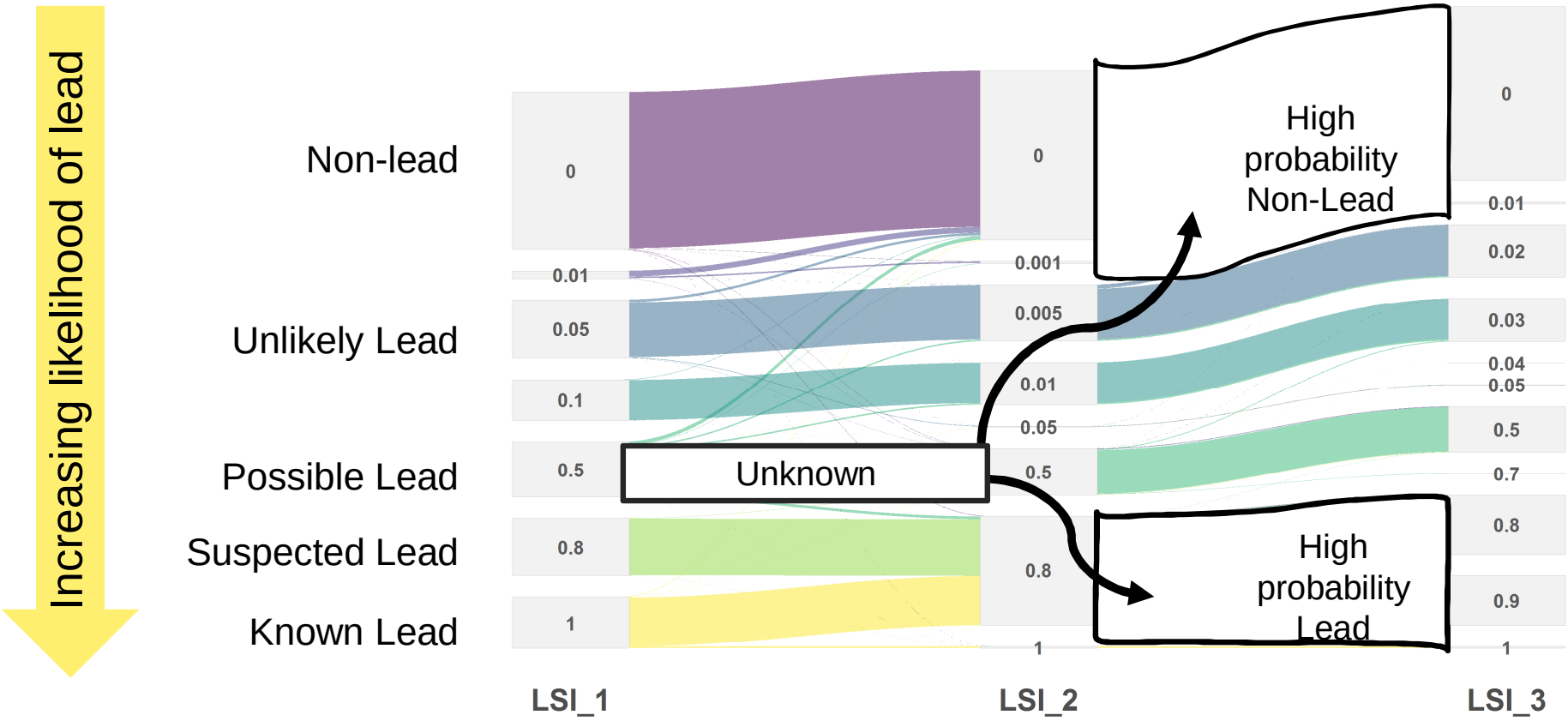
- Collect additional available data to train the model against initial data
  - Generate a “training” data set to improve the model output
  - Investigate up to 2000 properties/year
  - “score keeping by property”
- Run successive iterations with data to improve the certainty in the model

Examples of data used to train the model:

- ✓ **Records**
  - ✓ **microfiche data**
  - ✓ **asset data**
  - ✓ **watermain build records**
- ✓ **Water quality samples***
- ✓ **Pothole results***
- ✓ **LSL replacements***

***ongoing efforts**

Goal of predictive model over the life of a LSLR program and working at the extremes

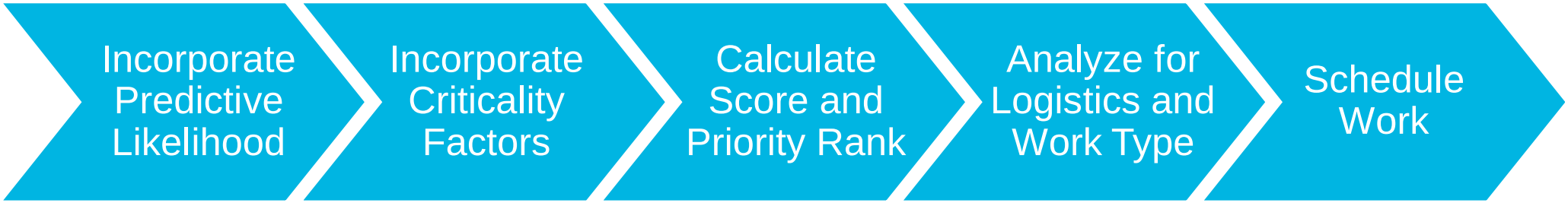


Going from model to results

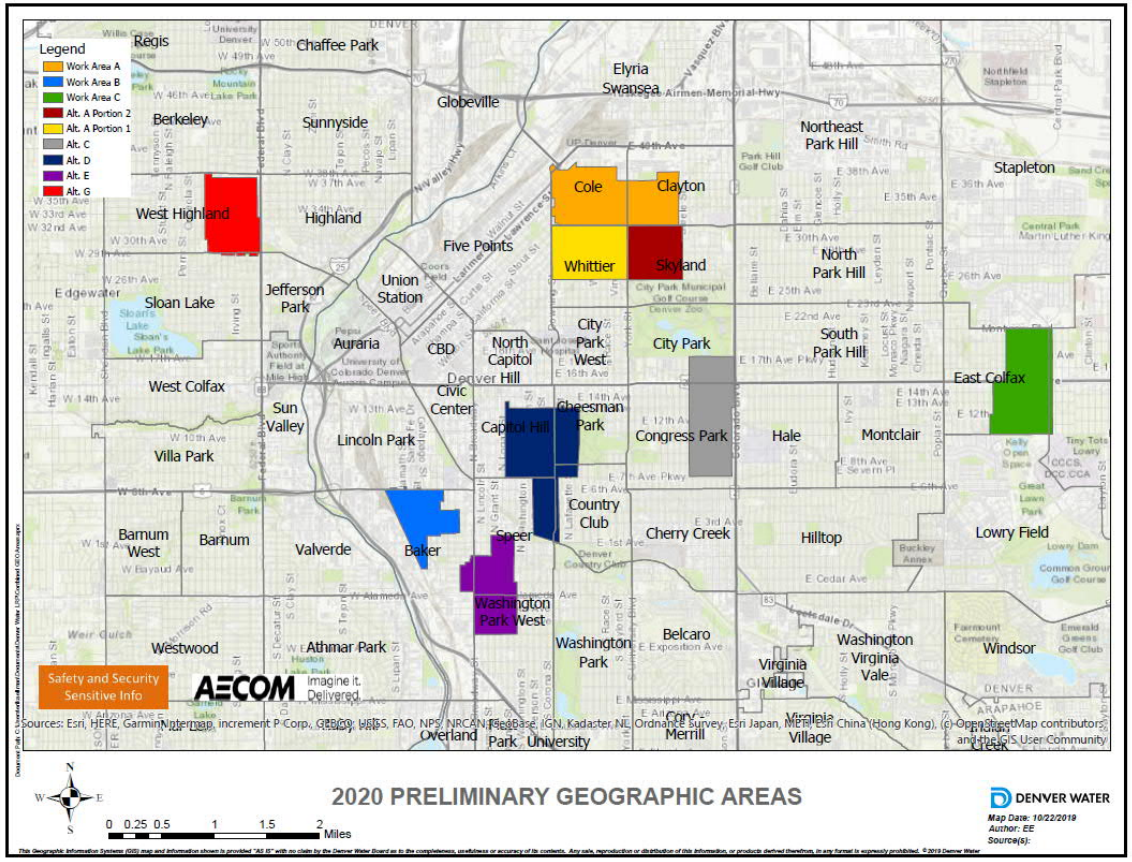
Applying the model to prioritize work

- Used the predictive model to prioritize:
  - Communications, outreach and education
  - Filter distribution
  - LSL replacement (geographic areas and individual replacements)
- Why did we do this?
  - Identify those most at risk to lead exposure
  - Coordinate planning with other infrastructure projects
  - Identify opportunities for construction efficiency
- How did we use the predictive model to do this?
  - Layer demographic and socioeconomic data on the LSL inventory model
    - Look for properties or areas with a high risk for lead in drinking water
    - Look for properties or areas with a high consequence from lead exposure

Sequencing tasks from planning to implementation



Candidate areas for LSL replacement in 2020



2020 Proposed Geographic ALSLR

DRAFT

DENVER WATER AECOM Imagine it. Delivered.

Denver Water Lead Reduction Program

October 22, 2019

AREA ID

2020 ALSLR—Geographic Work Area C

DESCRIPTION

Neighborhood (s): East Colfax
Boundaries: West/East: Tamarac St to Alley between Yosemite St and Akron St
North/South: Montview Blvd to E 11th Ave

Geographic Work Area C

Table for Geographic Work Area C

Lead Service Line Inventory			Census Data [1]					CDPHE	
Total Number of Taps	Total Number of Taps with Known, Suspected, & Possible Lead Services	Weighted Estimated Number of Lead Services	Population Estimates of Children Under 5 Years of Age	Population Estimates of Adults Above 65 Years of Age	Population Estimates of Women Between the Age of 15-50	Probability of Being Below the Federal Poverty Level [2]	Approximate Median Household Income	Low Income/ Minority	Average Number of eBLL Cases per Square Mile[3,4]
1114	977	713	439	250	1960	23.3%	$35,000	Y/Y	13.6

DATA DESCRIPTIONS

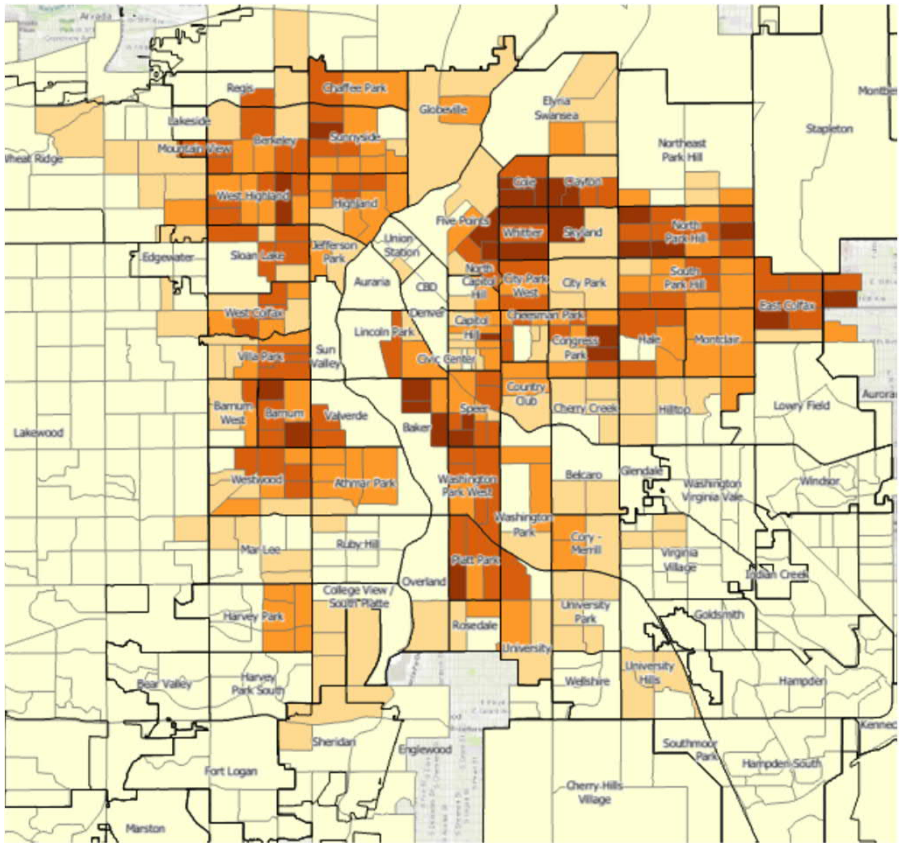
- Total Number of Taps: All taps within the lead service line inventory inside the work area.
- Total Number of Taps with Known, Suspected, & Possible Lead Services: All taps within the work area that have been identified as a known, suspected, or possible lead services.
- Expected Number of Lead Services: Total number of lead taps we expect to find based on the likelihood (probability for suspected (50% to 80%) and possible (80% to 90%) lead services) of finding lead.
- Population Estimates of Children Under 5 Years of Age: The number of children under the age of 5 living within the area's boundary.[1]
- Population Estimates of Adults Above 65 Years of Age: The number of adults above the age of 65 living within the area's boundary.[1]
- Population Estimates of Women Between the Age of 15-50: The number of women between the age of 15-50 living within the area's boundary.[1]
- Probability of Being Below the Federal Poverty Level: Percent chance of a household in the area's boundary being under the Federal Poverty Level.[2]
- Approximate Median Household Income: The household amount that divides the income distribution into two equal groups, half having income above that amount, and half having income below that amount within the area's boundary.
- Low Income/Minority: Is the area of concern considered an area of low income? (Yes or No). / Does the area of concern have a large minority population? (Yes or No).
- Average Number of elevated Blood Lead Level (eBLL) Cases per Square Mile: The average number of cases in a square mile where BLL is 5 micrograms per deciliter (ug/dL) or above.

DATA REFERENCES

1. 2010 Census data, CDPHE data sets.
2. 2013-2017 American Community Survey Estimates, CDPHE data sets.
3. Odds Ratio (OR) Contours from the Spatial Confounder-Adjusted Spatial Risk Model (Berg, et al, 2017).
4. Spatial Surveillance of Childhood Lead Exposure in a Targeted Screening State: An Application of Generalized Additive Models in Denver, Colorado, 2017

Prioritize filter distribution

- Different challenge for prioritization
- More properties need to be reached over a shorter time
- Used the predictive model with additional datasets to address the risk to public health from lead exposure



Who should use a predictive model?

When does predictive modelling make sense?

- How many service connections? And how much is (un)known about lead use?
- May not be efficient for small systems or low number of LSLs
- How scalable is this?
  - What is the duration of program (optimizing a 15 year program vs 1 year program with no data)
  - Are there enough data to train the model?
  - Are resources available to support data handling and management?
- Other drivers?
  - Transparency and messaging
  - Health equity and prioritization

Frequently asked questions

- When is the inventory perfect?
- How do you vet assumptions used to build the predictive model?
- Can we use the model to find:
  - Partial lead services remaining on the private side
  - What about galvanized (tricky!)
- When is this approach worthwhile?

Model to better understand the inventory
and/or
Model to plan and prioritize and message actions

Denver Water’s predictive and prioritization model

- LSL database with statistical tools to plan LSL replacement based on machine learning predictive model
  - Integration of multiple data sources to find the lead
- Risk-based prioritization modelling that considers both likelihood of lead and the consequence or risk impacts to the community
- An overall implementation plan that optimizes construction efficiency, field data validation, data integration, lead service inventory, and program implementation

Using the predictive and prioritization model to realize public health benefits sooner for all customers

Thank you

Quirien.Muylwyk@aecom.com