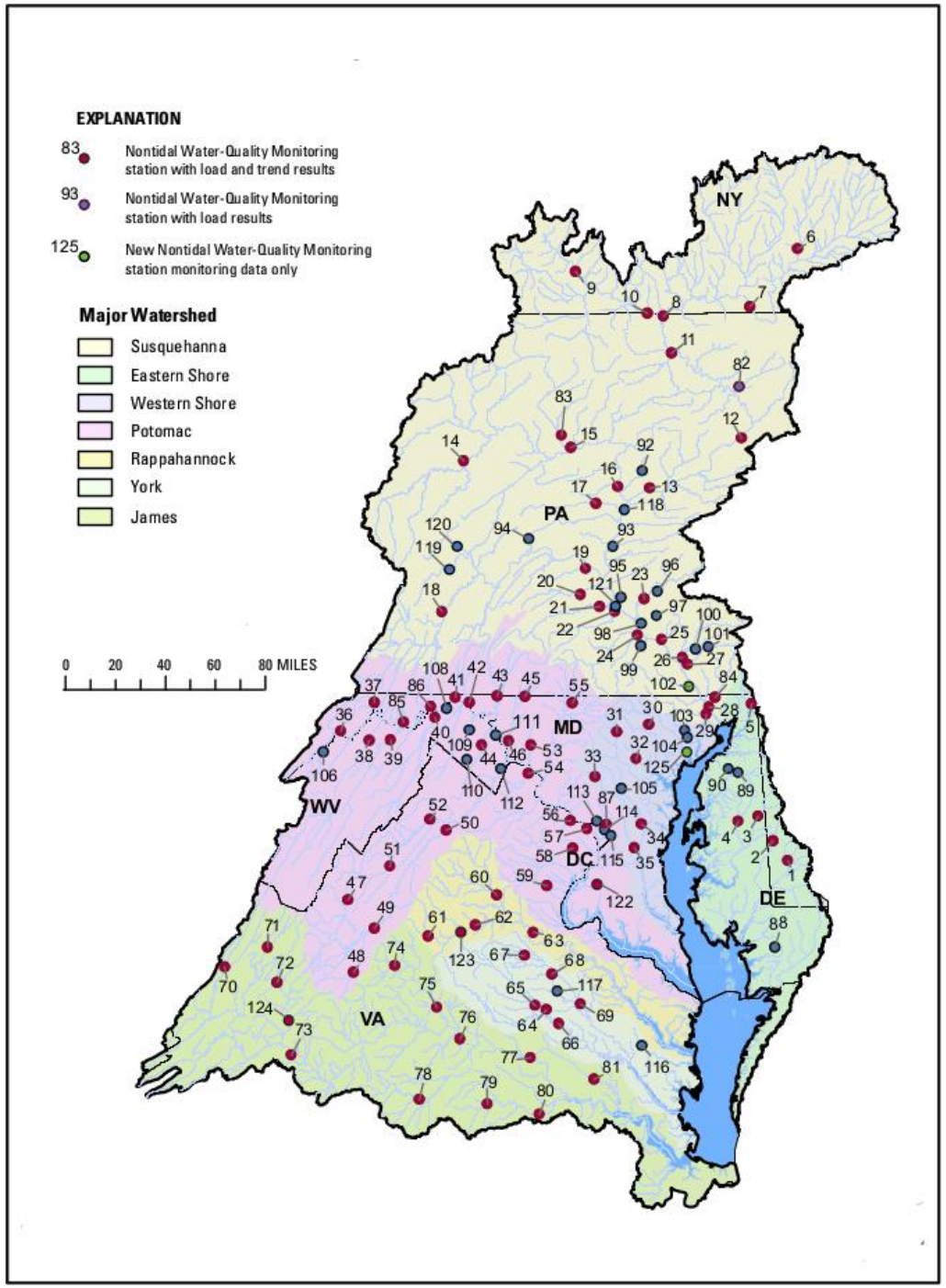
Identification of Monitoring Gaps in the Chesapeake Bay Nontidal Monitoring Network

Qian Zhang, UMCES
Matt Cashman, USGS

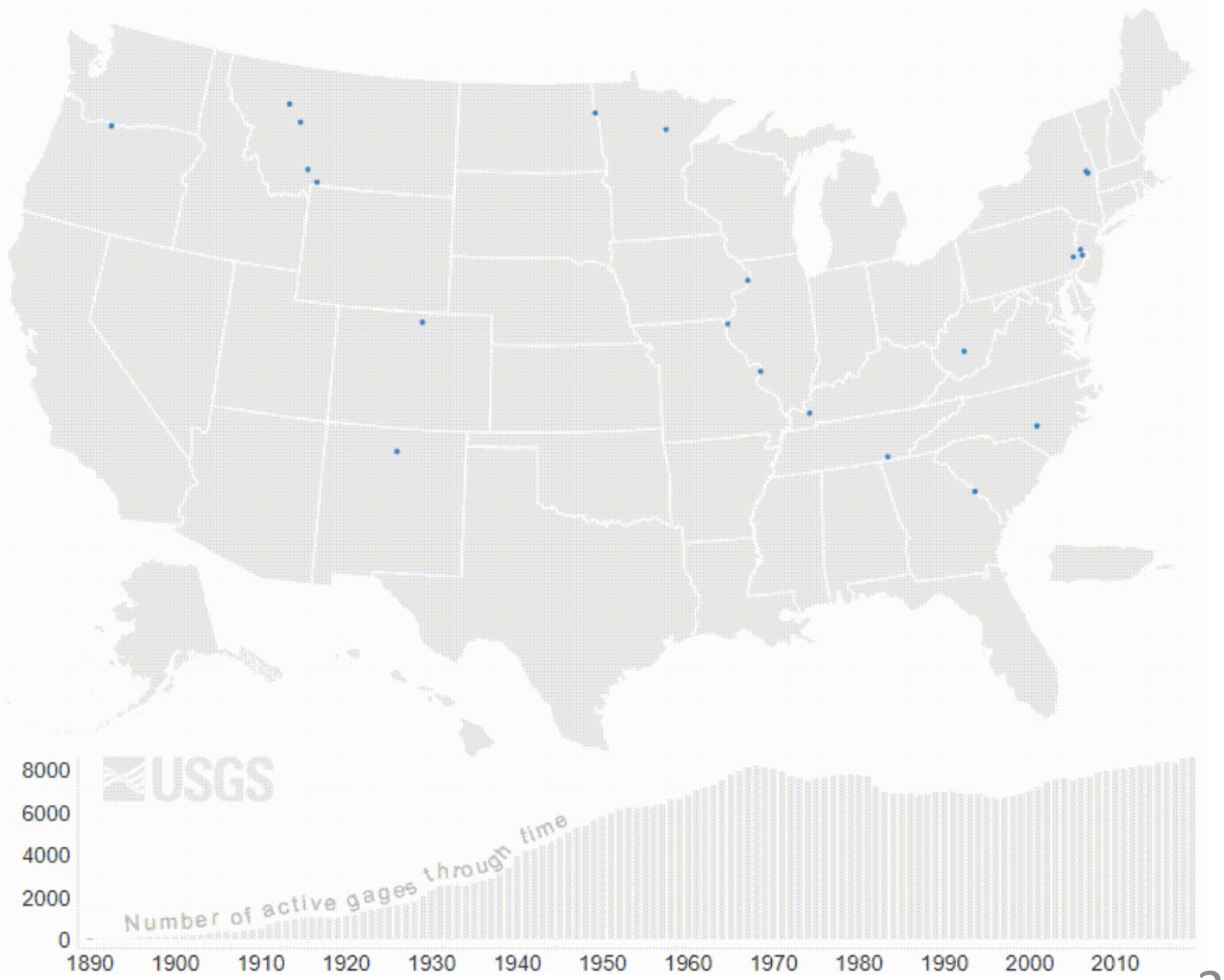


USGS National Streamgage Network

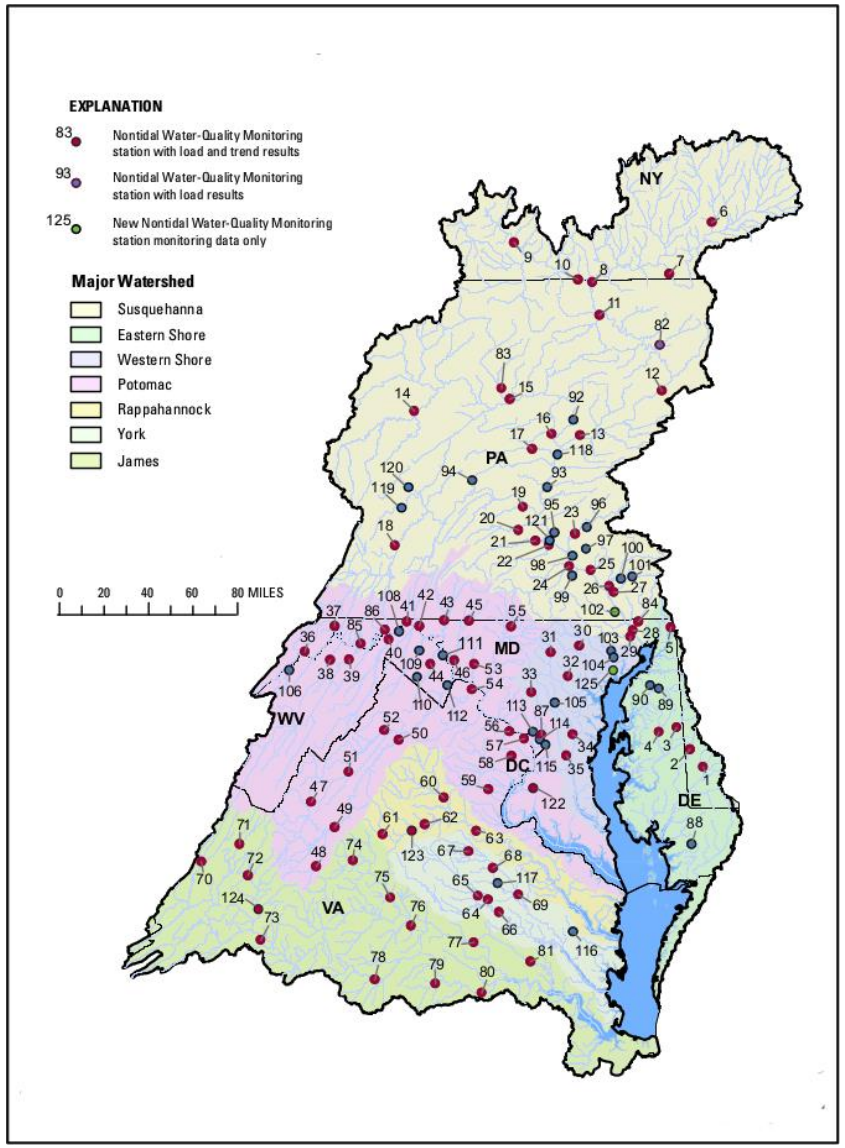
A network of networks

A collection of gages

An evolving source of water data



The CBNTN Monitoring Network (AOI)



Gaps in Streamgage Networks

- Are real
- Can be identified
- Must be understood and accounted for in models and analyses
- Monitoring can be optimized with an understanding of gaps

Gap Analysis

Identify gaps in the current monitoring network relative to important geographic variables using cumulative distribution analysis

Identify candidate sites to fill in those gaps

Conduct field reconnaissance and establish new monitoring locations

Gap Analysis

Graphical representation of relative redundance or inadequacy

Customized code applicable to different dimensions (i.e., water-quality parameters, physiographic provinces, or watershed sizes)

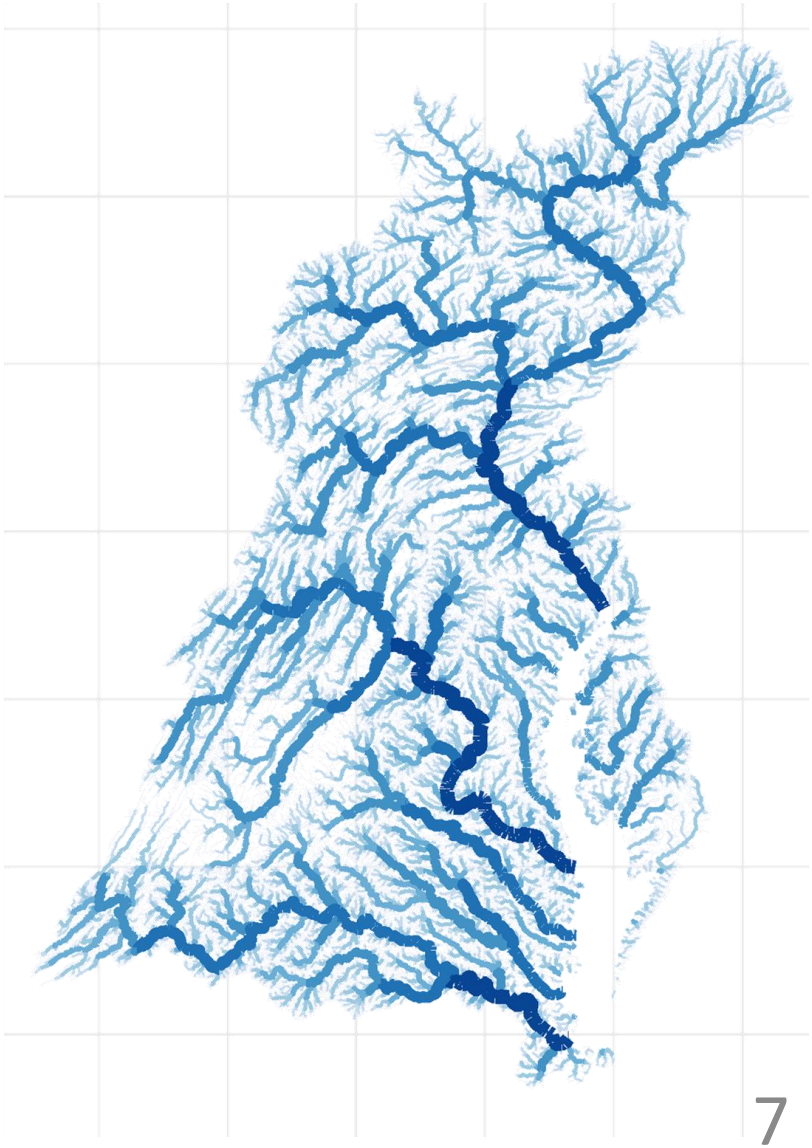
A pilot application implemented on the Delaware River Basin by Joel Blomquist and others

Gap Analysis

A basic question: “Are the samples representative of the population?”

- But challenges:
  - Unwieldy “big” data (large sizes, different sources)
  - Large GIS preprocessing *before* any analysis
  - Analysis methods (“how”)
  - Method reproducibility and iterative changes

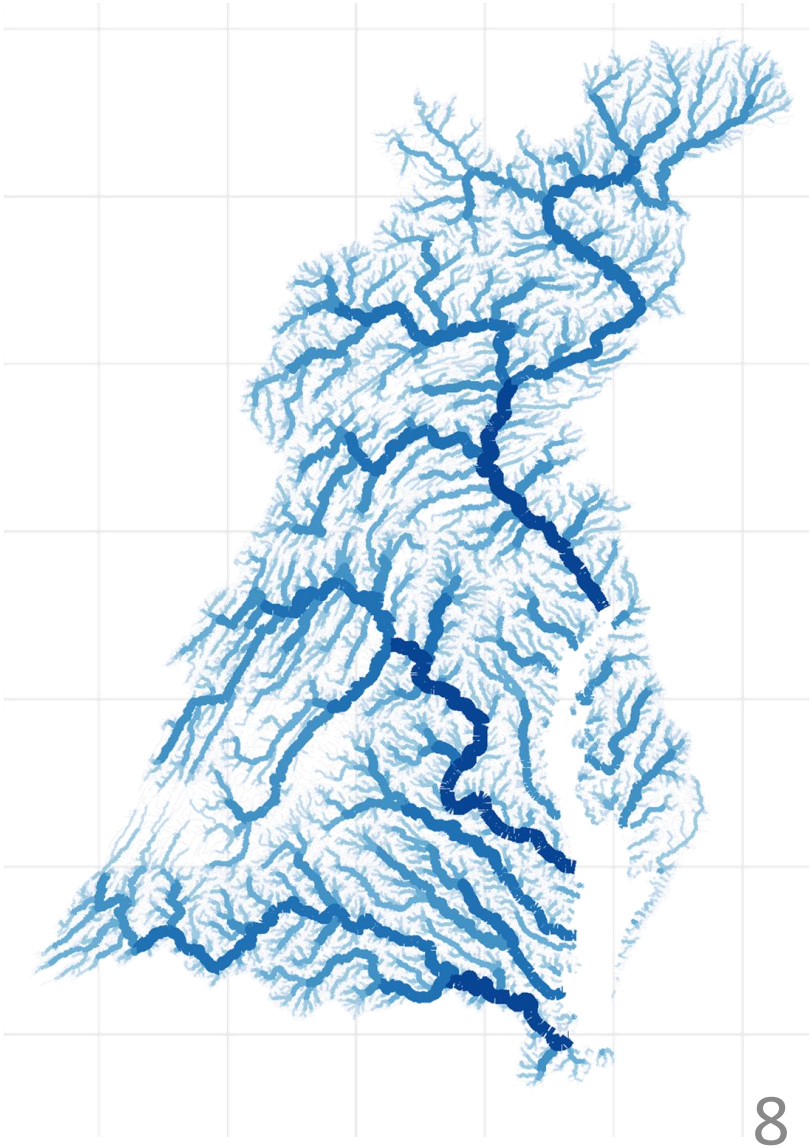
“Representative of ***what?***”



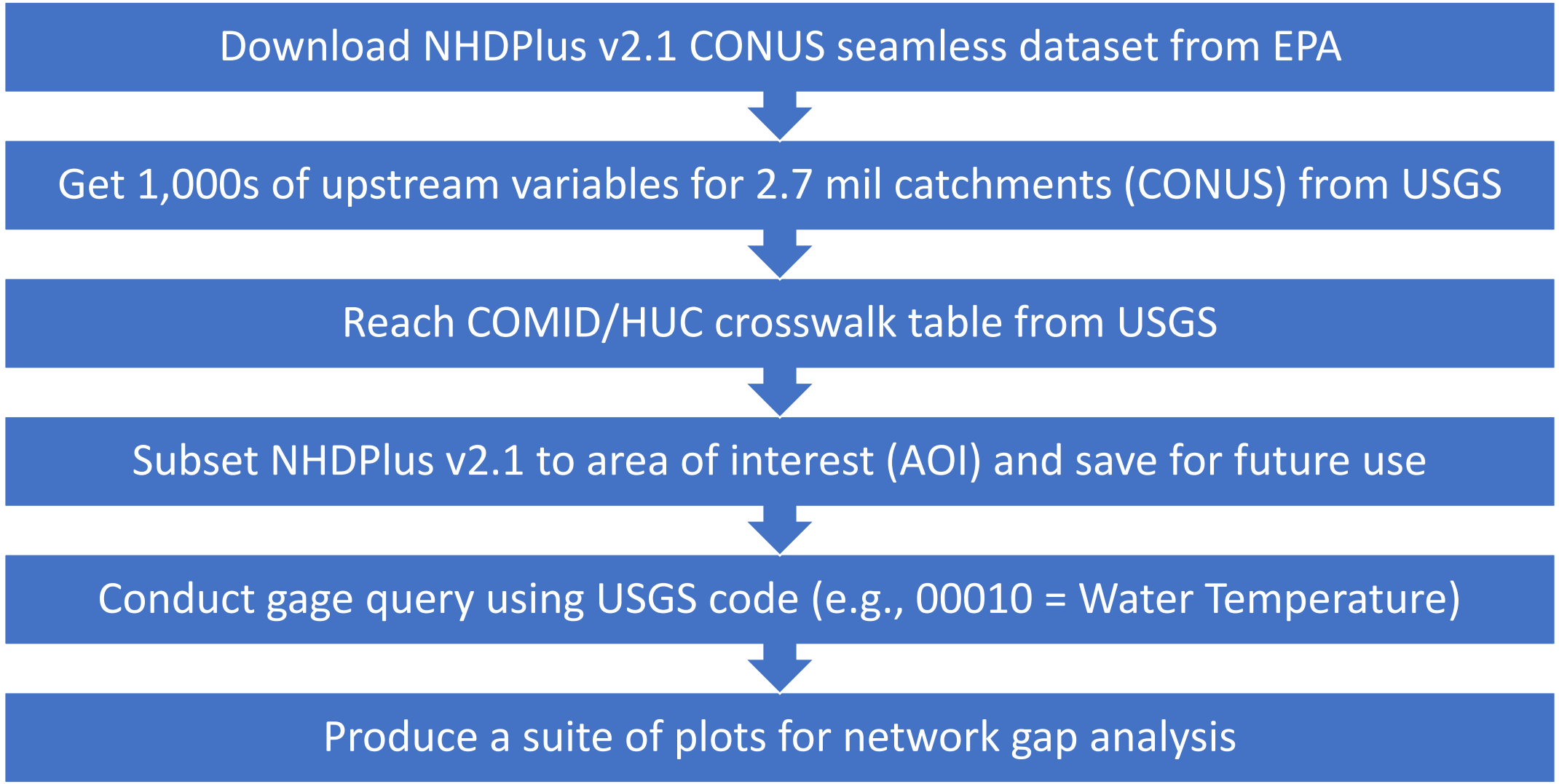
Gap Analysis

- A comparison of the representativeness of a monitoring network (“sample” or “target”) versus the underlying population it is supposed to represent (e.g., the river network)

- Useful for:
  - Predictive modeling
  - Identifying monitoring gaps to fill
  - Minimizing the impacts of monitoring cuts
  - Others…?



Tools to Streamline Network Analysis

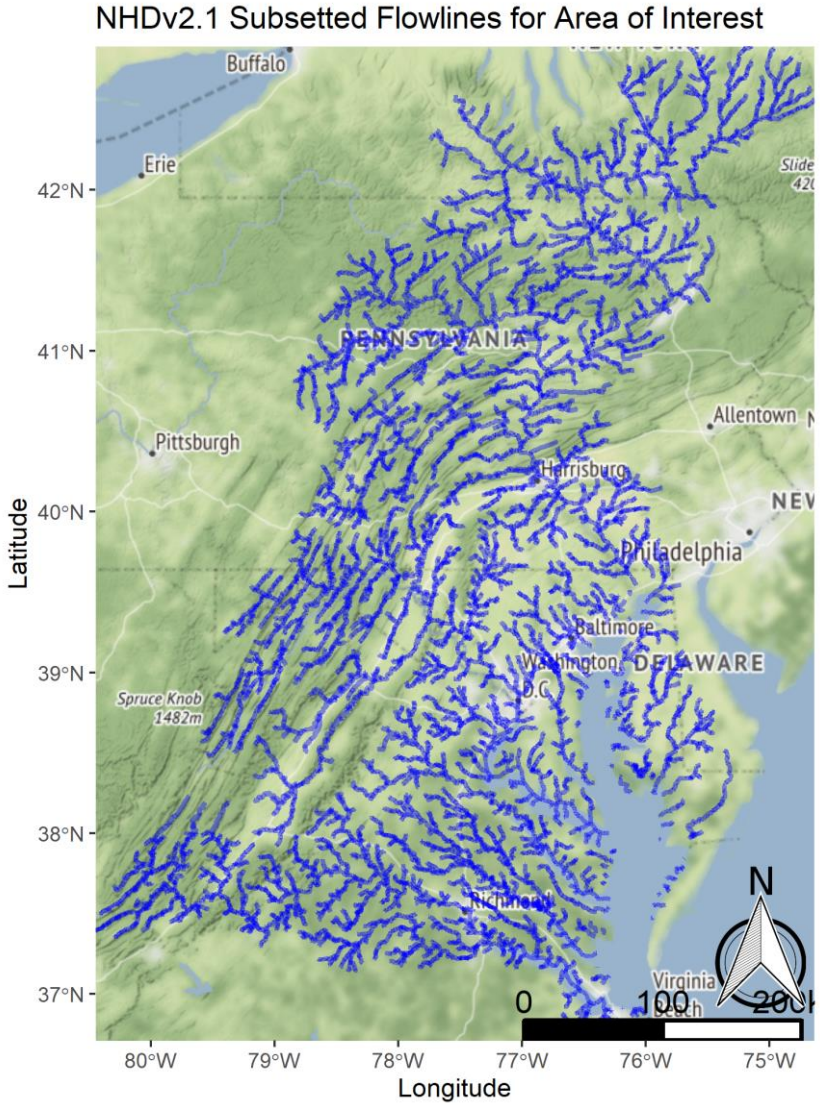


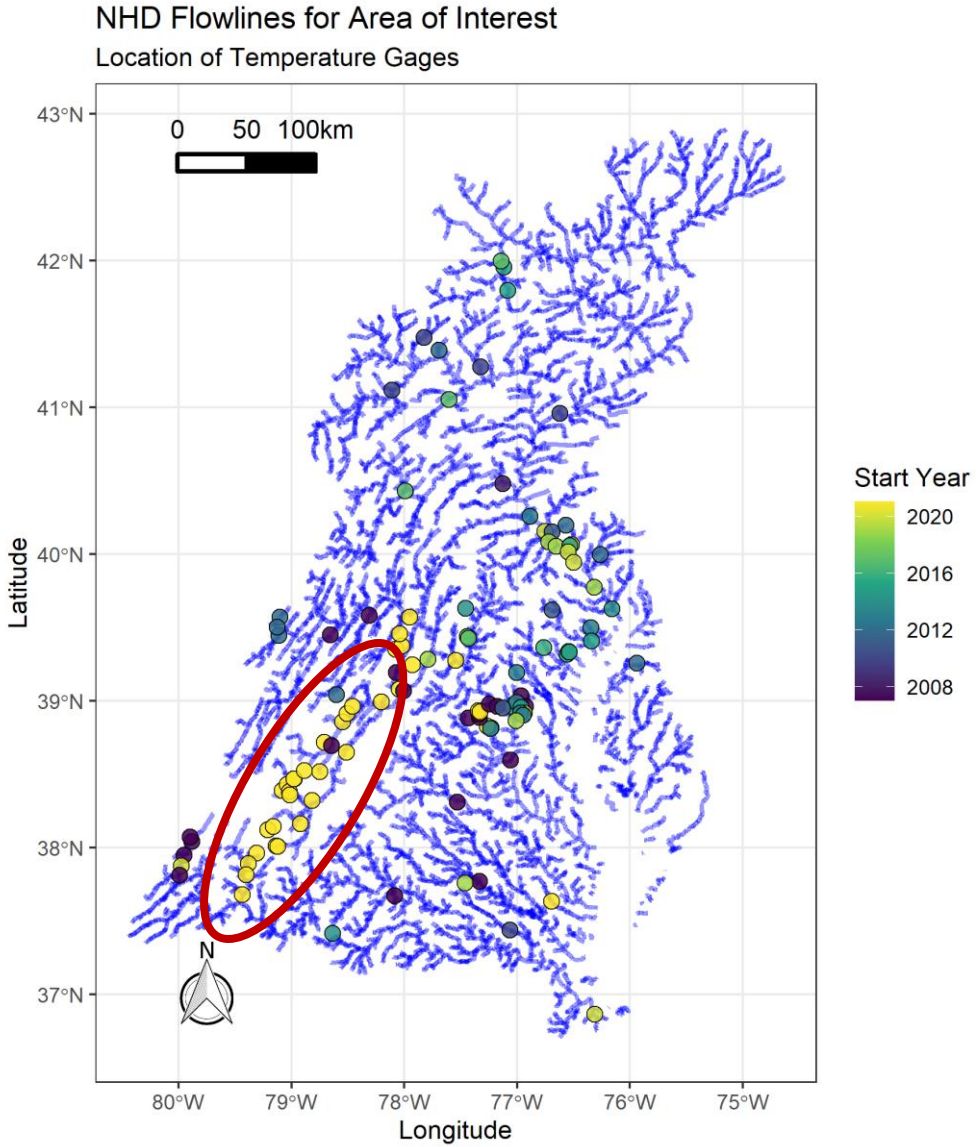
Landscape Queries

1. **Best Management Practices -** characteristics such as agricultural management practices and land in conservation practices.
2. **Chemical -** characteristics such as nitrogen application or toxicity weighted use.
3. **Climate and Water Balance Model -** characteristics such as model outputs of runoff, actual evapotranspiration or ground water storage.
4. **Climate -**characteristics such as mean precipitation, temperature, relative humidity, or evapotranspiration.
5. **Geology -** characteristics such as Hunt or Soller surficial geologies.
6. **Hydrologic -** characteristics such as base flow or infiltration excess overland flow. Hydrologic Modifications, characteristics such as dam storage or tile drains.
7. **Hydrologic Modifications -** characteristics such as dam storage or tile drains.
8. **Landscape -** characteristics such as NLCD, CDL or NWALT.
9. **Population Infrastructure -** characteristics such as population, housing, and road densities.
10. **Regions -** characteristics such as EcoRegions, Physiography or Hydrologic Landscapes.
11. **Soils -** characteristics such as STATSGO, soil salinity, and soil restrictive layer.
12. **Topographic Characteristics -** characteristics such as basin area, slope and elevation.
13. **Water use -** characteristics such as estimated freshwater withdrawls and estimated freshwater consumption by thermo-electric power plants

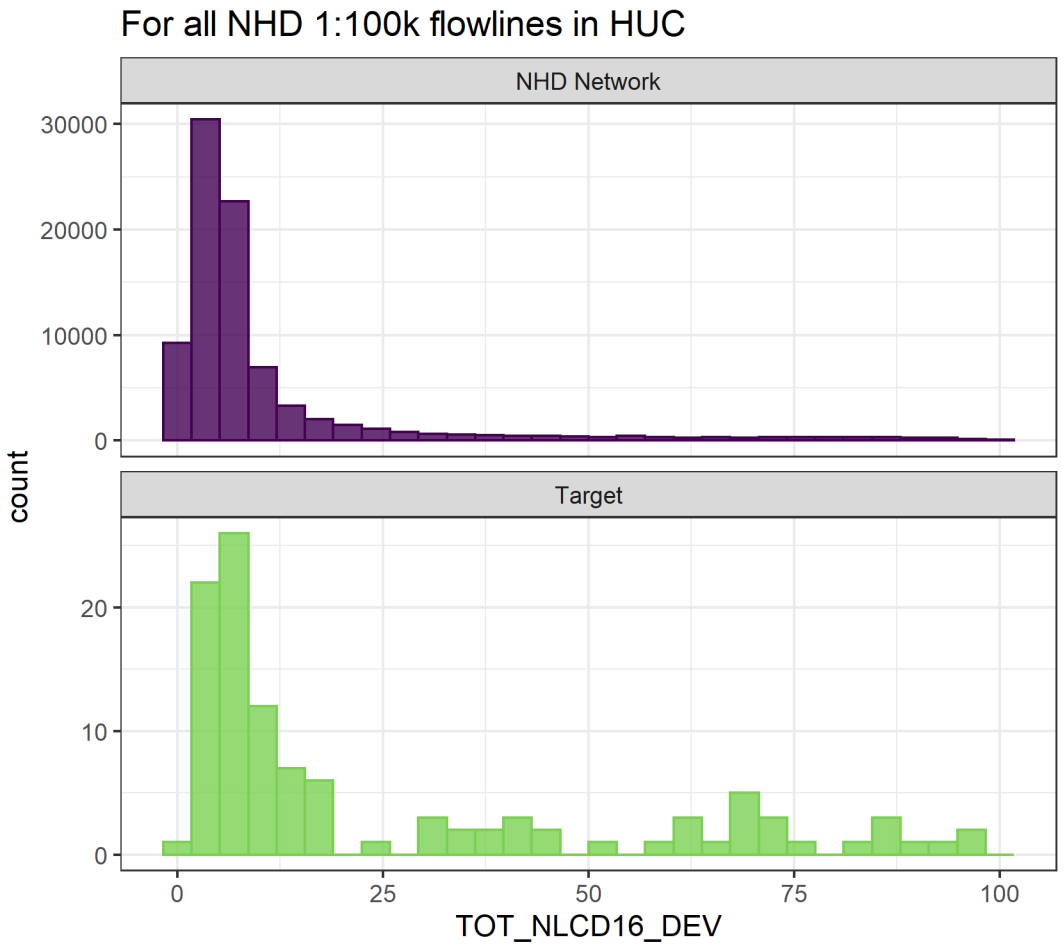
Wieczorek, M.E., Jackson, S.E., and Schwarz, G.E., 2018, Select Attributes for NHDPlus Version 2.1 Reach Catchments and Modified Network Routed Upstream Watersheds for the Conterminous United States (ver. 3.0, January 2021): U.S. Geological Survey data release, https://doi.org/10.5066/F7765D7V. https://www.sciencebase.gov/catalog/item/5669a79ee4b08895842a1d47

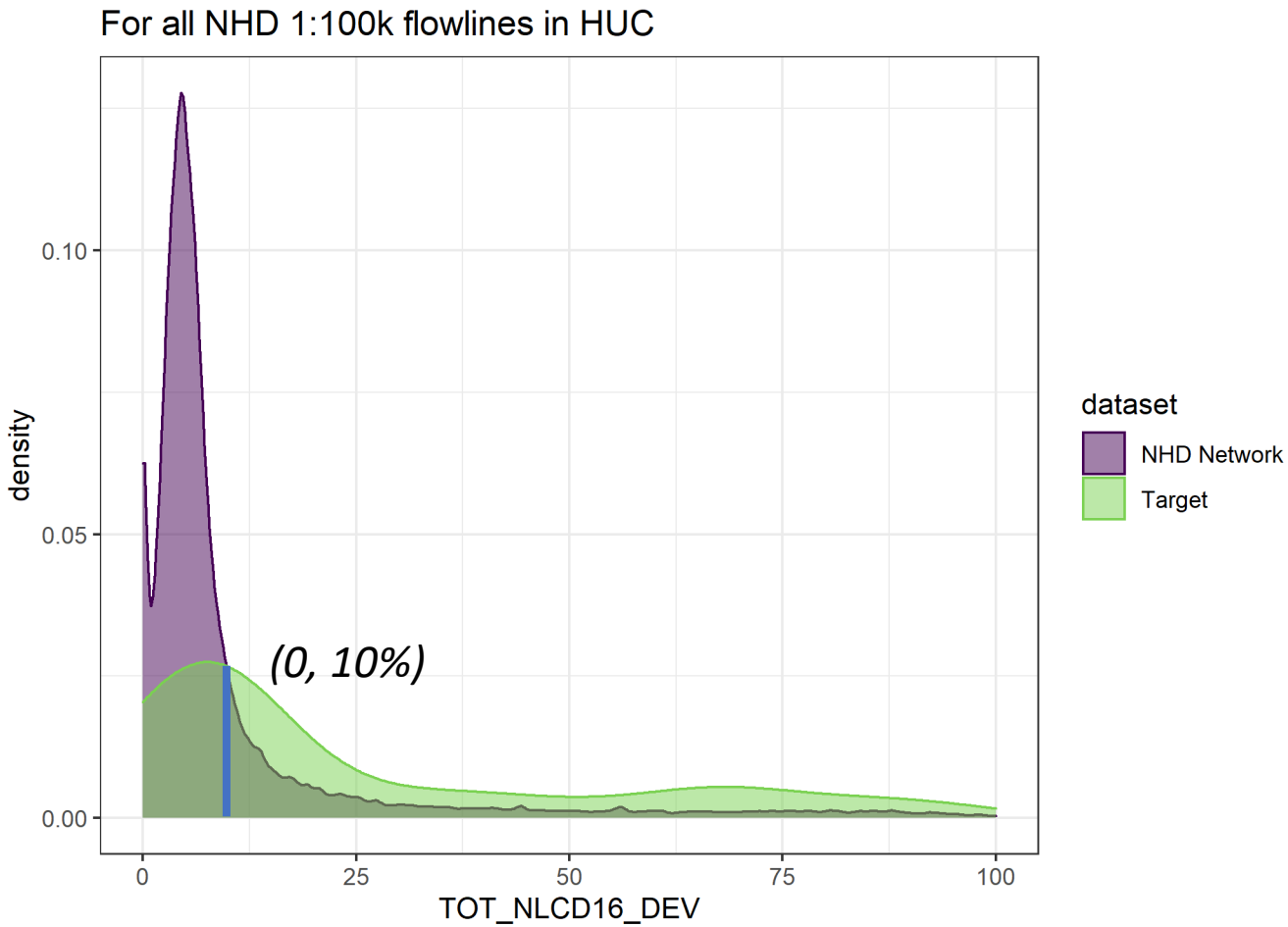
Flowline: Temperature (00100)



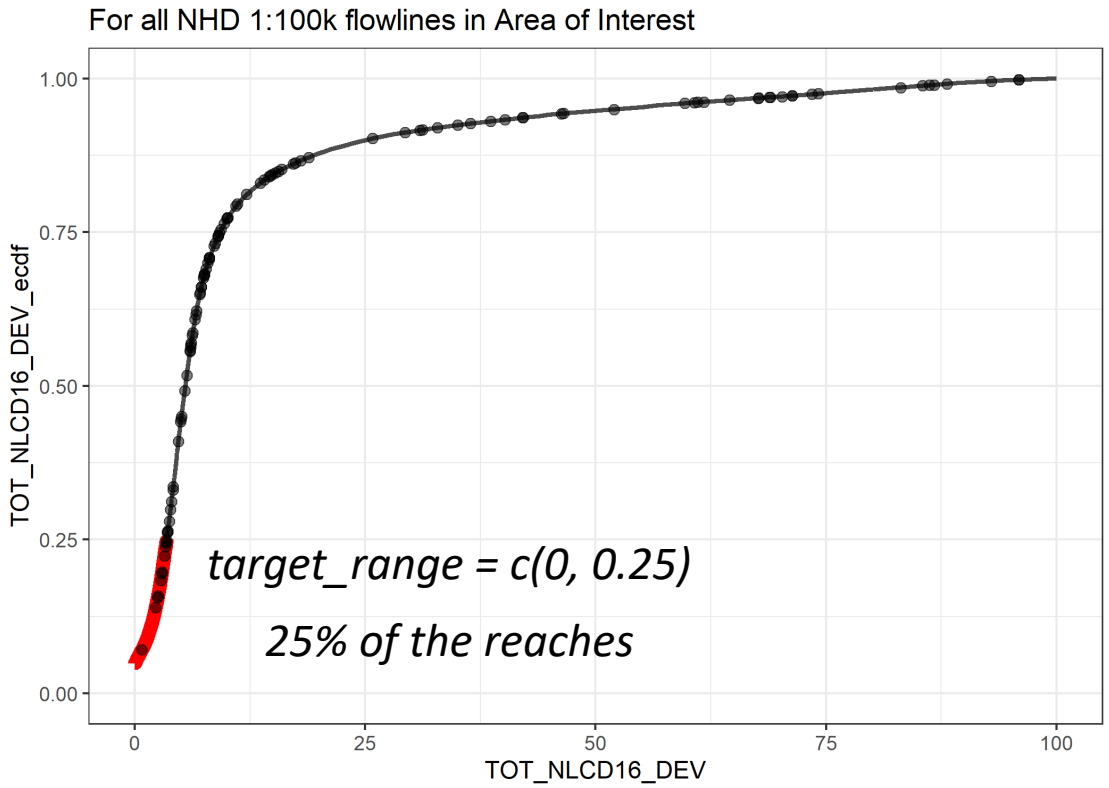


Histogram Analysis: Temperature

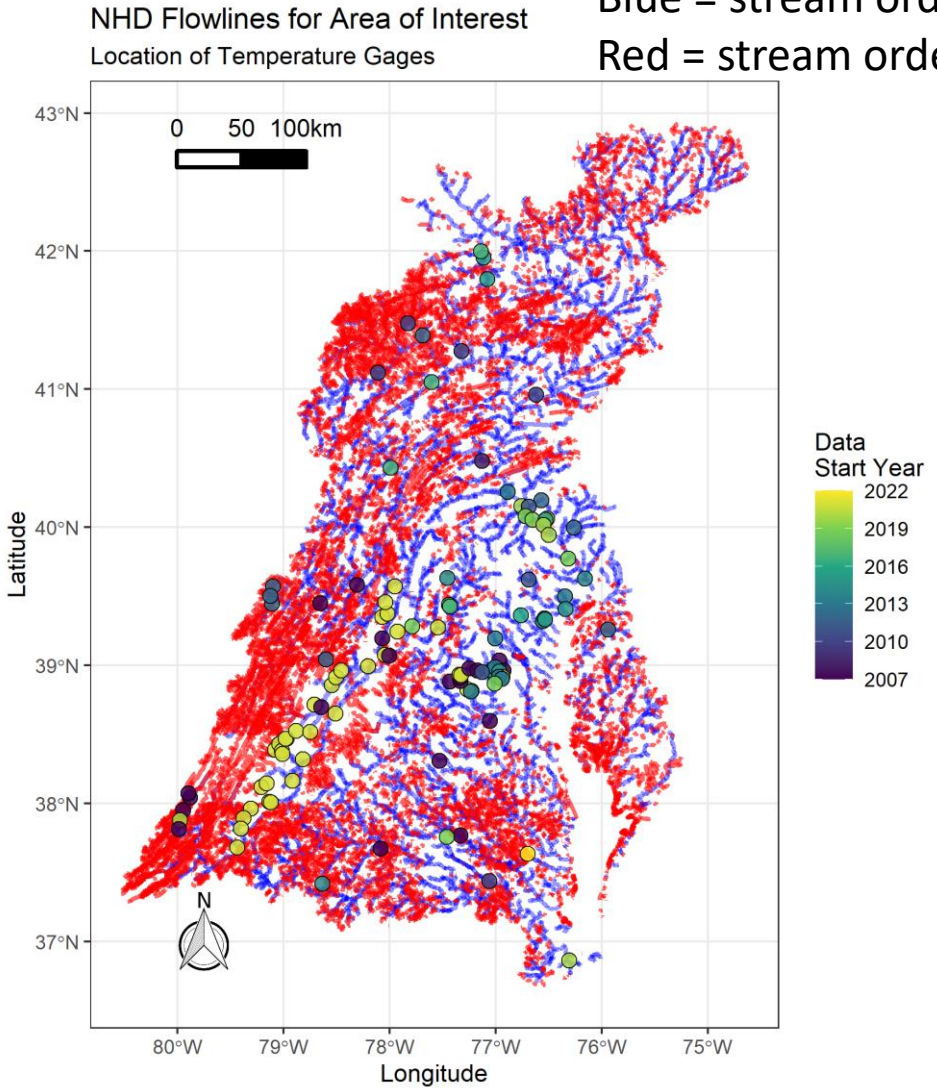




Histogram Analysis: Temperature



Blue = stream order 3
Red = stream order 5

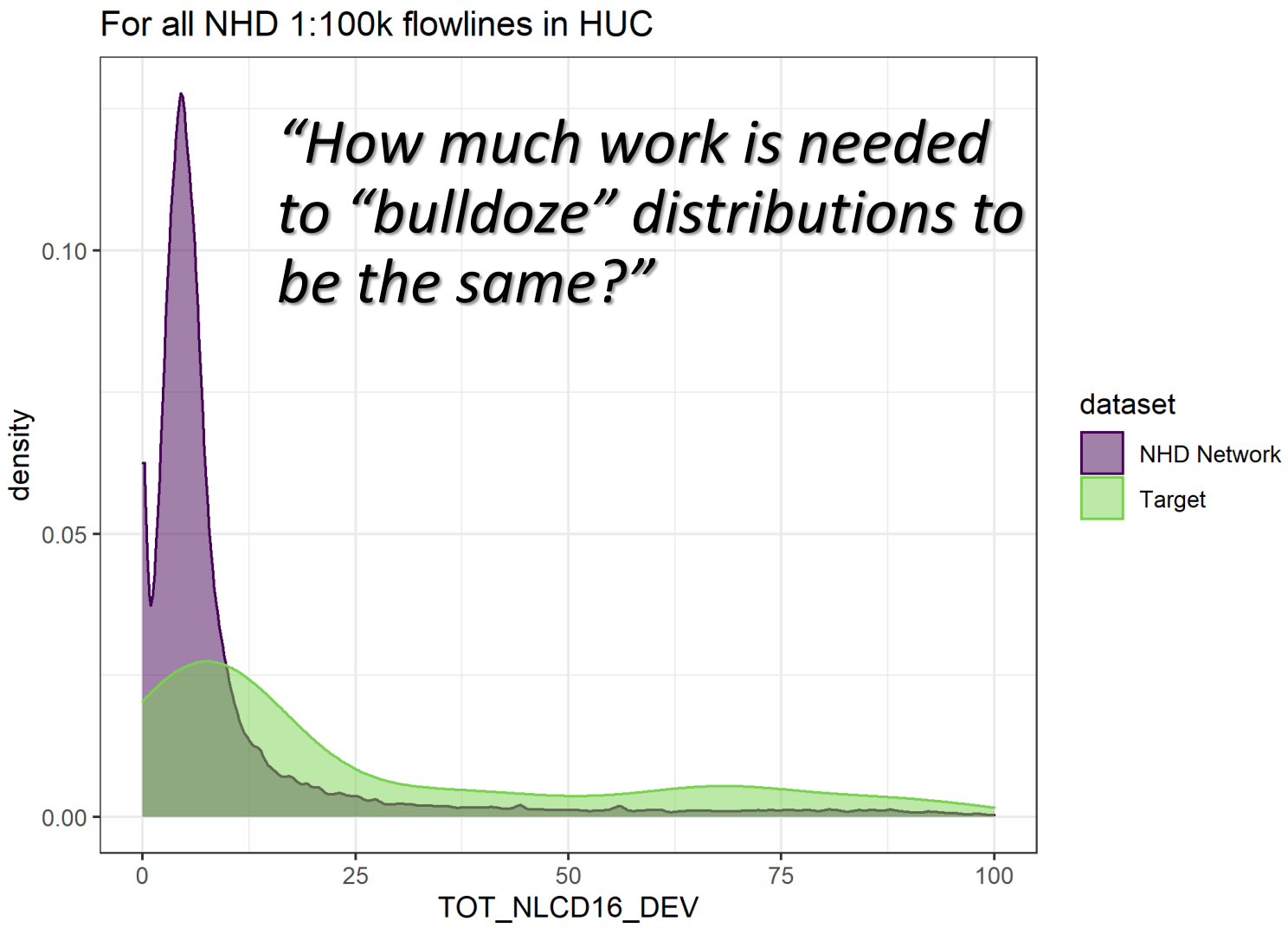


Gage Bias Quantification: Temperature

The **earth mover's distance** (**EMD**):

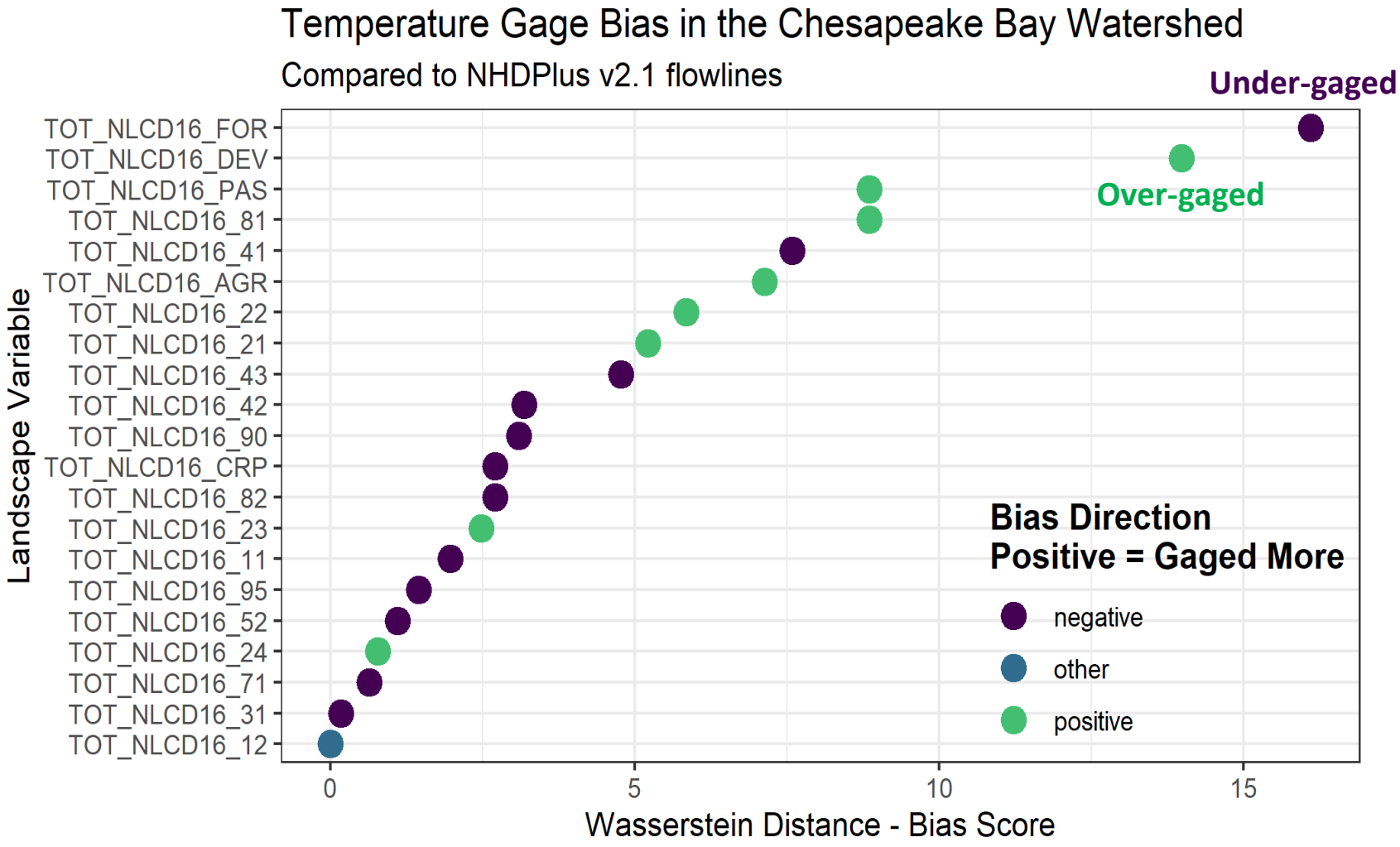
- Also known as the Wasserstein metric.
- A measure of the distance between two probability distributions over a region D.
- Informally, if the distributions are interpreted as two different ways of piling up a certain amount of earth (dirt) over the region D, the EMD is the minimum cost of turning one pile into the other; where the cost is assumed to be the amount of dirt moved times the distance by which it is moved.

Gage Bias Quantification: Temperature

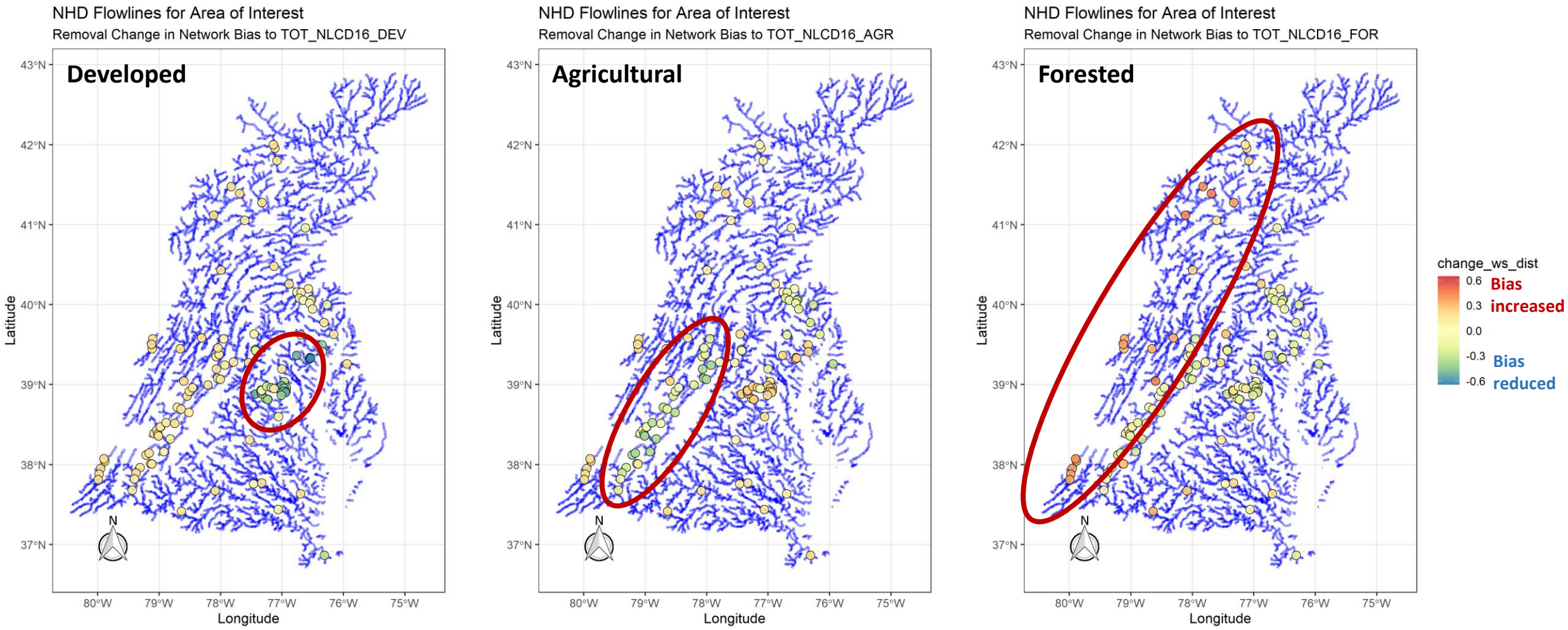


Gage Bias Quantification: Temperature

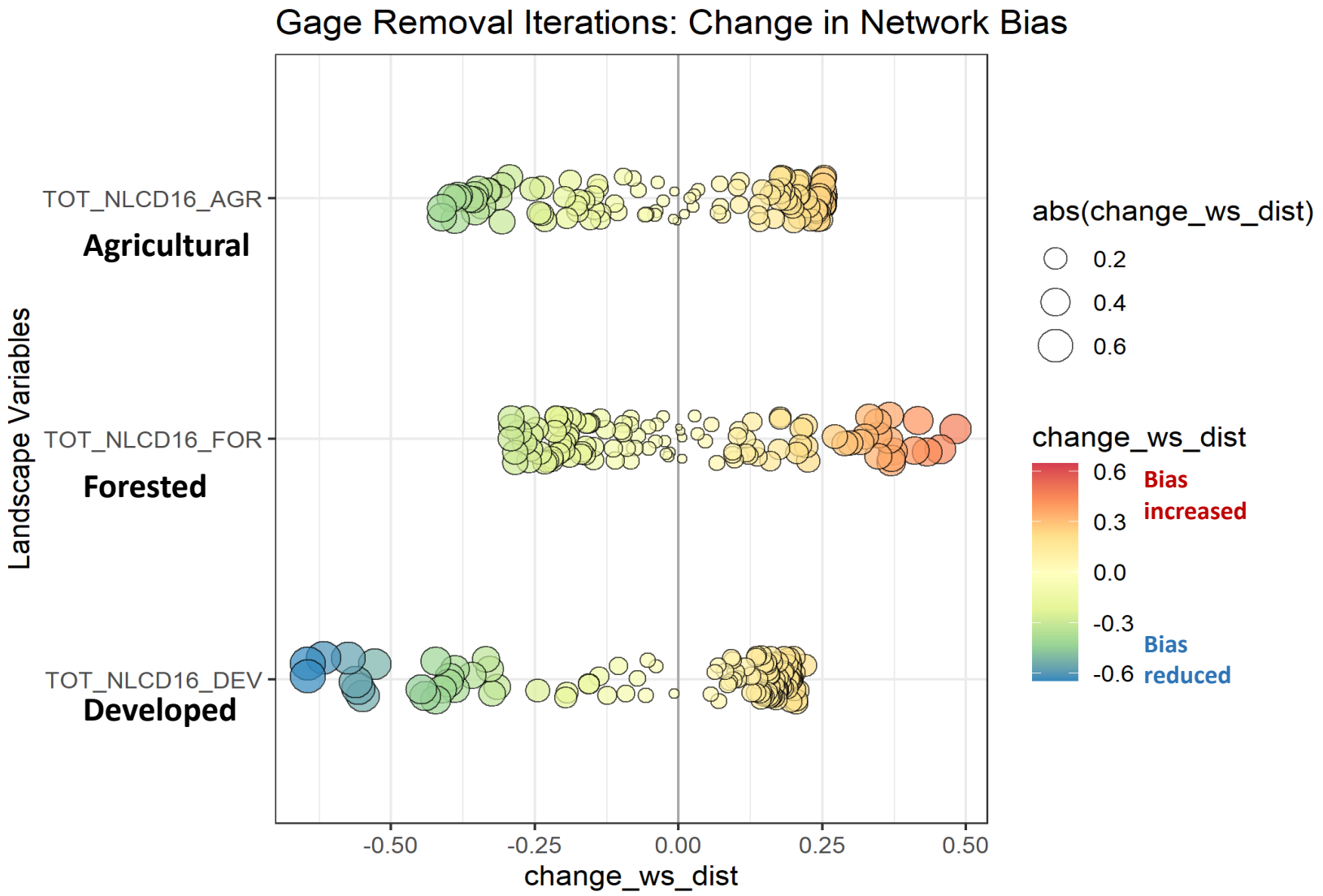
- Remove gage from network
- Re-calculate bias
- Calculate change in bias from baseline
- Iterate for each gage in network



Gage Removal Analysis: Temperature



Gage Removal Analysis: Temperature



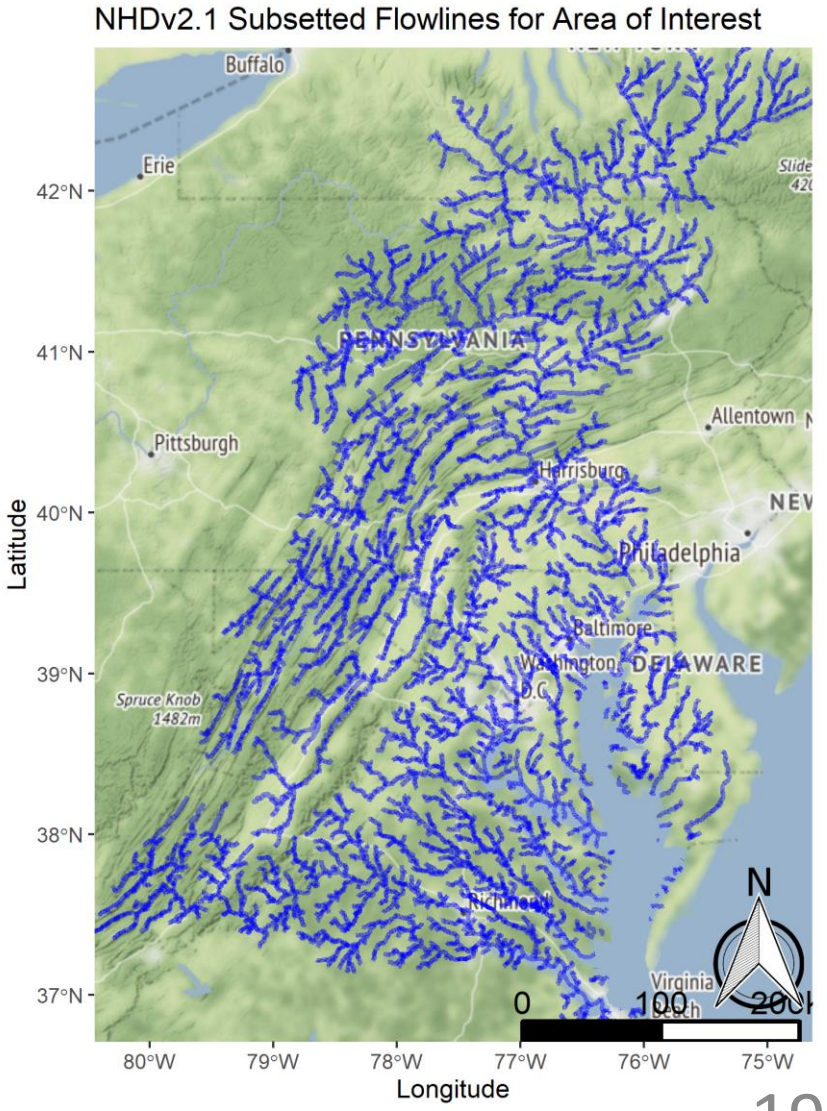
Future Work

Discrete Sampling

- Nutrients
- Sediment
- Bacteria
- PFAS and other contaminants
- Biological community

Continuous Sampling

- Temperature
- Conductivity
- Turbidity



Tools are only Tools

- These tools make network analyses “easier”
  - Handle “big” data
  - Avoid GIS pre-processing
  - Increase reproducibility
  - Built-in analysis methods

- Yet cannot answer the fundamental issue:
  - Is the network representative?
    - *Representative of “what”?*
  - What are your monitoring goals?